
W P E R L H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 06:17:41 1998; MasPar time 326.16 Seconds
Tabular output not generated. 543.181 Million cell updates/sec

Title: >US-08-951-733-19
Description: (1-3798) from US08951733.seq
Perfect Score: 3798
N.A. Sequence: 1 CCACGGTCGGCGGCGCT.....GGAATAGTCATCCCTGAT 3798
Comp: GGTGCGAGGCGCGTCCGA.....CCTTATCAGTAGGGACATA

Scoring table: TABLE default
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 88822 seqs, 23323279 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-issued
1:5_COMB 2:PCT9_COMB 3:backfiles1

Statistics: Mean 9.574; Variance 5.701; scale 1.679

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query % | Length | ID | Description | Pred. No. |
|------------|-------|---------|--------|----|------------------------------------|-----------|
| 1 | 67 | 1.8 | 7218 | 1 | US-08-232- Sequence 14, Applicati | 1.46e-25 |
| 2 | 49 | 1.3 | 7218 | 1 | US-08-232- Sequence 14, Applicati | 1.85e-14 |
| 3 | 34 | 0.9 | 215 | 1 | US-08-238- Sequence 5, Applicatio | 7.21e-06 |
| 4 | 35 | 0.9 | 215 | 1 | US-08-238- Sequence 5, Applicatio | 2.06e-06 |
| 5 | 25 | 0.7 | 66 | 1 | US-08-471- Sequence 144, Applicat | 2.91e-01 |
| 6 | 25 | 0.7 | 68 | 1 | US-07-971- Sequence 243, Applicat | 2.91e-01 |
| 7 | 25 | 0.7 | 69 | 1 | US-08-471- Sequence 142, Applicat | 2.91e-01 |
| 8 | 26 | 0.7 | 74 | 2 | PCT-US95-1 Sequence 94, Applicati | 9.62e-02 |
| 9 | 25 | 0.7 | 74 | 2 | PCT-US95-1 Sequence 100, Applicati | 2.91e-01 |
| 10 | 26 | 0.7 | 74 | 2 | PCT-US95-1 Sequence 100, Applicati | 9.62e-02 |
| 11 | 26 | 0.7 | 74 | 2 | PCT-US95-1 Sequence 99, Applicati | 2.91e-01 |
| 12 | 25 | 0.7 | 75 | 2 | PCT-US95-1 Sequence 99, Applicati | 2.91e-01 |
| 13 | 25 | 0.7 | 75 | 2 | PCT-US95-1 Sequence 99, Applicati | 2.91e-01 |
| 14 | 26 | 0.7 | 81 | 2 | PCT-US95-1 Sequence 92, Applicati | 9.62e-02 |
| 15 | 25 | 0.7 | 81 | 2 | PCT-US95-1 Sequence 98, Applicati | 2.91e-01 |
| 16 | 26 | 0.7 | 81 | 2 | PCT-US95-1 Sequence 98, Applicati | 9.62e-02 |
| 17 | 26 | 0.7 | 81 | 2 | PCT-US95-1 Sequence 92, Applicati | 9.62e-02 |
| 18 | 26 | 0.7 | 82 | 2 | PCT-US95-1 Sequence 97, Applicati | 9.62e-02 |
| 19 | 26 | 0.7 | 82 | 2 | PCT-US95-1 Sequence 97, Applicati | 9.62e-02 |

| | | | | | | |
|----|----|-----|------|---|-----------------------------------|----------|
| 20 | 25 | 0.7 | 242 | 1 | US-08-273- Sequence 1, Applicatio | 2.91e-01 |
| c | 21 | 0.7 | 1004 | 2 | PCT-US95-0 Sequence 7, Applicatio | 2.91e-01 |
| c | 22 | 0.7 | 1386 | 2 | PCT-US95-0 Sequence 2, Applicatio | 2.91e-01 |
| c | 23 | 0.7 | 1611 | 2 | PCT-US93-0 Sequence 3, Applicatio | 2.91e-01 |
| c | 24 | 0.6 | 65 | 1 | US-08-471- Sequence 145, Applicat | 2.49e+00 |
| c | 25 | 0.6 | 65 | 1 | US-08-471- Sequence 145, Applicat | 8.62e-01 |
| c | 26 | 0.6 | 66 | 2 | PCT-US95-1 Sequence 93, Applicati | 8.62e-01 |
| c | 27 | 0.6 | 66 | 1 | US-08-471- Sequence 144, Applicat | 8.62e-01 |
| c | 28 | 0.6 | 68 | 1 | US-08-471- Sequence 143, Applicat | 2.49e+00 |
| c | 29 | 0.6 | 68 | 1 | US-08-471- Sequence 143, Applicat | 8.62e-01 |
| c | 30 | 0.6 | 69 | 1 | US-08-471- Sequence 142, Applicat | 8.62e-01 |
| c | 31 | 0.6 | 84 | 1 | US-08-300- Sequence 25, Applicati | 2.49e+00 |
| c | 32 | 0.6 | 84 | 2 | PCT-US94-0 Sequence 25, Applicati | 2.49e+00 |
| c | 33 | 0.6 | 84 | 1 | US-08-209- Sequence 23, Applicati | 2.49e+00 |
| c | 34 | 0.6 | 84 | 1 | US-08-133- Sequence 120, Applicat | 2.49e+00 |
| c | 35 | 0.6 | 108 | 1 | US-08-357- Sequence 2, Applicatio | 2.49e+00 |
| c | 36 | 0.6 | 225 | 1 | US-08-299- Sequence 18, Applicati | 2.49e+00 |
| c | 37 | 0.6 | 225 | 1 | US-07-807- Sequence 15, Applicati | 2.49e+00 |
| c | 38 | 0.6 | 242 | 1 | US-08-273- Sequence 1, Applicatio | 8.62e-01 |
| c | 39 | 0.6 | 1404 | 3 | 5480796-8 Patent No. 5480796. | 2.49e+00 |
| c | 40 | 0.6 | 1640 | 1 | US-07-807- Sequence 11, Applicati | 2.49e+00 |
| c | 41 | 0.6 | 1640 | 1 | US-08-299- Sequence 11, Applicati | 2.49e+00 |
| c | 42 | 0.6 | 2253 | 3 | 5457037-2 Patent No. 5457037. | 2.49e+00 |
| c | 43 | 0.6 | 2454 | 1 | US-07-872- Sequence 1, Applicatio | 8.62e-01 |
| c | 44 | 0.6 | 2625 | 3 | 5457037-4 Patent No. 5457037. | 2.49e+00 |
| c | 45 | 0.6 | 3336 | 3 | 5457037-1 Patent No. 5457037. | 2.49e+00 |

ALIGNMENTS

RESULT 1
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC xxxxxx

Sequence 14, Application US/08232463
Sequence 14, Application US/08232463
Patent No. 5670367

GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/995,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELEPHONE: (703)836-9300
TELEFAX: (703)836-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:


```

CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc.feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
CC SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 0.9% Score 35; DB 1; Length 215;
Best Local Similarity 13.28; Pred.No. 2.06e-06;
Matches 26; Conservative 76; Mismatches 95; Indels 0; Gaps 0;

Db 6 MSSSVVSRASCNDRKAKDKGNTTSSWTTDCCNRTGWGCDTTTYRVNNDSGHNKYSAN 65
Cp 1616 AGTCCGCGCAGCTCATCTTCCAGCTGAGCTCCTGCAGGAGAGCTTGGCATGCTTCCCCA 1557
Db 66 YNYGNNVGAAKHYTHYTHNVSGADSKYVDSDVNSAGTSSSSNGTGDNRSGADSYSSKT 125
Cp 1556 GGGAGATGAATCTTCTTGCTGCTCCTGAGGAAGCGGCGTTCTGTTGTGCTGGAGCCCGAGA 1497
Db 126 AMTSRNTGTANNVDSRNMGDAVSGSDKNTKKHAKNSADGKVGSKNNGDRNNRYGTGT 185
Cp 1496 GGCCTGGGGGACAGCGCGCGGAGCGGAGCGCCGACGAAGCCGTACACCTGCCAGGGGC 1437
Db 186 KSNVSNNGGGRKRDVS 202
Cp 1436 TGCTGTGCTGCGGGAGC 1420

RESULT 5
ID US-08-471-052A-144 STANDARD; DNA; UNC; 66 BP.
XXXXXX
DT
DE Sequence 144, Application US/08471052A
CC Sequence 144, Application US/08471052A
CC Patent No. 5625033
CC GENERAL INFORMATION:
CC APPLICANT: Kay, B. K.
CC APPLICANT: Fowlkes, D. M.
CC TITLE OF INVENTION: Totally Synthetic Affinity Reagents
CC NUMBER OF SEQUENCES: 166
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennlie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/471,052A
CC FILING DATE: 06-JUNE-1995
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872

```


[illegible]

CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 81 BP; 6 A; 6 C; 4 G; 5 T; 60 OTHER.

Query Match 0.7%; Score 25; DB 2; Length 81;
Best Local Similarity 7.7%; Pred. No. 2.91e-01;
Matches 5; Conservative 20; Mismatches 40; Indels 0; Gaps 0;

Db 13 VNNYNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 72
Qy 946 CCATCCACCACATCCGTGGCCGCCAGCACCGGGGGCCCCCATCCACATCGCGCC 1005
Db 73 ACCAC 77
Qy 1006 ACCAC 1010

Search completed: Fri Dec 25 06:26:06 1998
Job time : 505 secs.

MPERLH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 24 07:42:45 1998; MasPar time 55.38 Seconds
Tabular output not generated. 780.716 Million cell updates/sec

Title: >US-08-951-733-20
Description: (1-1154) from US08951733.pep
Perfect Score: 8624
Sequence: 1 HASGQRCVLLRTWEALAPAT.....TALEAAANPALPSDFKTLTD 1154

Scoring table: PAM 150
Gap 11

Searched: 116738 segs, 37463448 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir58
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 54.465; Variance 114.185; scale 0.477

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | Pred. No. |
|------------|-------|-------------|--------|-------|------------------------------|-----------|
| 1 | 314 | 3.6 | 884 | 2 | S53396 telomerase catalytic | 1.90e-31 |
| 2 | 162 | 1.9 | 660 | 1 | QBEB3 BHLF1 protein - human | 4.72e-07 |
| 3 | 153 | 1.8 | 504 | 2 | JC1306 virion protein homolo | 8.51e-06 |
| 4 | 136 | 1.6 | 240 | 2 | B24264 proline-rich protein | 1.59e-03 |
| 5 | 141 | 1.6 | 276 | 2 | B38955 hypothetical protein | 1.53e-04 |
| 6 | 142 | 1.6 | 310 | 1 | PIHUSD salivary proline-rich | 2.61e-04 |
| 7 | 136 | 1.6 | 316 | 2 | S16681 homeotic protein - hu | 1.59e-03 |
| 8 | 137 | 1.6 | 628 | 2 | S01955 hypothetical protein, | 1.18e-03 |
| 9 | 129 | 1.5 | 264 | 2 | ORF4 protein - Orf vi | 1.24e-02 |
| 10 | 131 | 1.5 | 300 | 2 | D34768 proline-rich protein | 6.92e-03 |
| 11 | 126 | 1.5 | 302 | 2 | S11790 nodulation protein no | 2.92e-02 |
| 12 | 129 | 1.5 | 309 | 2 | S10889 proline-rich protein | 1.24e-02 |
| 13 | 132 | 1.5 | 317 | 2 | A28986 proline-rich protein | 5.17e-03 |
| 14 | 127 | 1.5 | 566 | 2 | S22933 testis-specific prote | 1.19e-02 |
| 15 | 132 | 1.5 | 924 | 2 | S27923 gene LF3 protein - hu | 5.17e-03 |
| 16 | 131 | 1.5 | 1106 | 2 | J00405 hypothetical 119.5K P | 6.92e-03 |
| 17 | 130 | 1.5 | 1236 | 2 | E70977 hypothetical protein | 9.26e-03 |
| 18 | 117 | 1.4 | 204 | 2 | A39066 proline-rich protein | 3.53e-01 |
| 19 | 125 | 1.4 | 227 | 2 | C29149 proline-rich protein | 3.87e-02 |
| 20 | 123 | 1.4 | 240 | 2 | A24264 proline-rich protein | 6.79e-02 |
| 21 | 123 | 1.4 | 256 | 2 | A50533 tumor-associated anti | 6.79e-02 |
| 22 | 118 | 1.4 | 322 | 1 | S00034 nodulation protein no | 2.69e-01 |
| 23 | 119 | 1.4 | 347 | 2 | S10571 epithelial tumor anti | 2.05e-01 |

| | | | | | | |
|----|-----|-----|------|---|------------------------------|----------|
| 24 | 121 | 1.4 | 358 | 1 | WMBE38 infected cell protein | 1.18e-01 |
| 25 | 125 | 1.4 | 381 | 2 | S16506 hypothetical protein | 3.87e-02 |
| 26 | 125 | 1.4 | 402 | 2 | A45056 osteogenic protein 2 | 3.87e-02 |
| 27 | 123 | 1.4 | 403 | 2 | S52796 prpL2 protein - human | 6.79e-02 |
| 28 | 118 | 1.4 | 431 | 2 | S09824 hypothetical protein | 2.69e-01 |
| 29 | 123 | 1.4 | 439 | 2 | S51939 chitinase (EC 3.2.1.1 | 6.79e-02 |
| 30 | 124 | 1.4 | 464 | 2 | S22697 extensin - Volvox car | 5.13e-02 |
| 31 | 119 | 1.4 | 515 | 2 | S10372 epithelial tumor anti | 2.05e-01 |
| 32 | 114 | 1.3 | 202 | 2 | B36795 hypothetical protein | 7.86e-01 |
| 33 | 115 | 1.3 | 260 | 2 | S22373 proline-rich protein | 6.03e-01 |
| 34 | 116 | 1.3 | 270 | 2 | S34361 miaE protein - Salmon | 4.62e-01 |
| 35 | 116 | 1.3 | 295 | 2 | B48013 proline-rich proteogl | 4.62e-01 |
| 36 | 114 | 1.3 | 301 | 2 | E29149 proline-rich protein | 7.86e-01 |
| 37 | 114 | 1.3 | 321 | 1 | A28663 nodulation protein no | 7.86e-01 |
| 38 | 115 | 1.3 | 338 | 2 | JC5707 HYA22 protein - human | 6.03e-01 |
| 39 | 114 | 1.3 | 373 | 2 | S54545 hypothetical protein | 7.86e-01 |
| 40 | 115 | 1.3 | 392 | 1 | PIHUB6 salivary proline-rich | 6.03e-01 |
| 41 | 114 | 1.3 | 408 | 2 | A56186 cyclin E - African cl | 7.86e-01 |
| 42 | 113 | 1.3 | 704 | 2 | A30411 synapsin Ia - rat | 1.02e-00 |
| 43 | 114 | 1.3 | 705 | 2 | A35363 synapsin I splice for | 7.86e-01 |
| 44 | 113 | 1.3 | 1255 | 2 | B35175 epistatin B - human | 1.02e-00 |
| 45 | 115 | 1.3 | 3149 | 1 | QBEB8 BPLF1 protein - human | 6.03e-01 |

ALIGNMENTS

RESULT 1 S53396 #type complete
ENTRY telomerase catalytic chain EST2 - yeast (Saccharomyces
TITLE cerevisiae)
ALTERNATE_NAMES protein L8543.12; protein YLR318W
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 05-May-1995 #sequence_revision 01-Sep-1995 #text_change
13-Sep-1998

ACCESSIONS S53396

REFERENCE S53390

#authors Du, Z.

#submission submitted to the EMBL Data Library, February 1995

#description The sequence of S. cerevisiae cosmid 8543.

#accession S53396

##molecule_type DNA

##residues 1-884 ##label DUZ

##cross-references EMBL:U20618; NID:92258165; PID:9662136; MIPS:YLR318W

##experimental_source strain S288C (AB972)

GENETICS

#gene

##cross-references SGD:S0004310; MIPS:YLR318W

#map_position 12R

SUMMARY #length 884 #molecular-weight 102662 #checksum 7604

Query Match 3.6%; Score 314; DB 2; Length 884;

Best Local Similarity 24.3%; Pred. No. 1.90e-31;

Matches 118; Conservative 122; Mismatches 202; Indels 44; Gaps 37;


```

sequence IS1222
#formal_name Enterobacter agglomerans
03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change
17-Jul-1998
ACCESSIONS
B38965; S42922
REFERENCE
#authors
Steibl, H.D.; Lewecke, F.M.
#journal
Gene (1995) 156:37-42
#title
IS1222: analysis and distribution of a new insertion sequence
in Enterobacter agglomerans 339.
#accession
B38965
#status
preliminary
#molecule_type DNA
#residues
1-276 #label STE
#cross-references GB:X78052; NID:g459246; PID:g459248
#experimental_source strain Kleeberger, 1983
GENETICS
#mobile_element insertion sequence IS1222
#length 276 #molecular_weight 31718 #checksum 3927
SUMMARY
Query Match 1.68; Score 141; DB 2; Length 276;
Best Local Similarity 33.78; Pred. No. 3.53e-04;
Matches 30; Conservative 23; Mismatches 30; Indels 6; Gaps 6;
Db 42 ITLALERR-FCYRR-IMOLLREGLVNHKRVYRLYHLSGLGVKRRR-RKGLATERL 98
QY 585 VTETFOKNRFLFYKRSWKSGLSIGRQHLKRV-QLRELSEAEVQHRREARPAULTSRL 643
Db 99 PLL-RPAAPNLTWSDDFVMDALATGRRIK 126
QY 644 RFPKPDGLRPVIMDYVGA-RFRRREK 671

RESULT 6
ENTRY
#authors
PIHUSD #type complete
#journal
salivary proline-rich glycoprotein precursor PRB4 (large
#title
allele) - human
CONTAINS
basic proline-rich protein IB-5; proline-rich peptide P-D
#formal_name Homo sapiens #common_name man
#length 1984 #sequence_revision 12-Apr-1996 #text_change
05-Sep-1997
ACCESSIONS
S03176; S03175; S10890; D25372; E38355; A03295; A61294;
S62891
REFERENCE
S02127
#authors
Lyons, K.M.; Stein, J.H.; Smithies, O.
#journal
Genetics (1988) 120:267-278
#title
Length polymorphisms in human proline-rich protein genes
generated by intragenic unequal crossing over.
#cross-references MUID:89121440
#accession
S03176
#status
translation not shown
#molecule_type DNA
#residues
35-310 #label LY1
#cross-references EMBL:X07715
#note
large allele
#accession
S03175
#status
translation not shown
#molecule_type DNA
#residues
35-36, 'P', '38-112, 155-310 #label LY2
#cross-references EMBL:X07704
#note
medium allele
#accession
S10890
#status
preliminary; translation not shown
#molecule_type DNA
#residues
1-38, 60-112, 'T', '114-115, 'P', '117-121, 185-271, 'A', 273-310
#label LY3
#cross-references EMBL:X07882; NID:g35647; PID:g296670
A92492
REFERENCE
#authors
Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.
#journal
J. Biol. Chem. (1985) 260:11123-11130
#title
Differential RNA splicing and post-translational cleavages in
the human salivary proline-rich protein gene system.
#cross-references MUID:85289325

```

```

#accession D25372
#molecule_type mRNA
#residues
1-36, 'E', '38-112, 'T', '114-115, 'P', '117-121, 185-271, 'A',
273-310 #label MAE
REFERENCE
A38355
#authors
Kauffman, D.L.; Bennick, A.; Blum, M.; Keller, P.J.
#journal
Biochemistry (1991) 30:3351-3356
#title
Basic proline-rich proteins from human parotid saliva:
relationships of the covalent structures of ten proteins
from a single individual.
#cross-references MUID:91190884
#accession
E38355
#molecule_type protein
#residues
241-254, 'KN', 257-310 #label KAU
REFERENCE
A03295
#authors
Saitoh, E.; Isemura, S.; Sanada, K.
#journal
J. Biochem. (1983) 93:495-502
#title
Complete amino acid sequence of a basic proline-rich peptide,
P-D, from human parotid saliva.
#cross-references MUID:83186122
#accession
A03295
#molecule_type protein
#residues
241-310 #label SAI
REFERENCE
A61294
#authors
Shimomura, H.; Kanai, Y.; Sanada, K.
#journal
J. Biochem. (1983) 93:857-863
#title
Amino acid sequences of glycopeptides obtained from basic
proline-rich glycoprotein of human parotid saliva.
#accession
A61294
#molecule_type protein
#residues
54-57, 'E', 59-73, 'R', 82-101 #label SHI
REFERENCE
S62891
#authors
Charlton, A.J.; Baxter, N.J.; Lilley, T.H.; Haslam, E.;
McDonald, C.J.; Williamson, M.P.
#journal
FEBS Lett. (1996) 382:289-292
#title
Tannin interactions with a full-length human salivary
proline-rich protein display a stronger affinity than with
single proline-rich repeats.
#accession
S62891
#molecule_type protein
#residues
241-252 #label CHA
#note
amino end of peptide designated basic proline-rich
protein IB-5
#note
it is unclear from the peptide sequence whether this is
a product of the PRB2 (PIR:PIHUPF) or PRB4 (this
entry) gene
GENETICS
#gene
GDB:PRB4
#cross-references GDB:119514; OMIM:180990
#map_position 12p13.2-12p13.2
#introns
22/1; 34/1
#note
the list of introns may be incomplete
CLASSIFICATION
#superfamily proline-rich protein
KEYWORDS
glycoprotein; saliva; tandem repeat
FEATURE
1-16
241-310
66,87,171
108,150,192,213,
234
#binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY
#length 310 #molecular_weight 31351 #checksum 3960
Query Match 1.68; Score 142; DB 1; Length 310;
Best Local Similarity 28.28; Pred. No. 2.61e-04;
Matches 40; Conservative 41; Mismatches 52; Indels 9; Gaps 9;
Db 168 OGNOSQO-PPPHGKXPERPP-OGG-NOSHRPPPP-CKPER-PPPOGNGSQO-P-PP 220
QY 199 QLGANTQARPPPHASGPRRLGCRANWNSVRAGVPLGLPAPGARRRGGSASRSLPLPK 258

```



```

#status preliminary
#molecule_type DNA
#residues 1-300 #label ROB
#cross-references GB:X58438; NID:g53181; PID:g53182
REFERENCE
S22570
#authors Roberts, S.G.E.; Layfield, R.; McDonald, C.J.
#journal Nucleic Acids Res. (1991) 19:5205-5211
#title The mouse proline-rich protein MP6 promoter binds
isoprenaline-inducible parotid nuclear proteins via a
highly conserved NFkB/rel-like site.
#cross-references MUID:92020206
#accession S22570
#molecule_type DNA
#residues 1-14 #label RO2
#cross-references EMBL:X61126
CLASSIFICATION #superfamily proline-rich protein
SUMMARY #length 300 #molecular-weight 31129 #checksum 8443

Query Match 1.5%; Score 131; DB 2; Length 300;
Best Local Similarity 25.2%; Pred. No. 6.92e-03;
Matches 34; Conservative 40; Mismatches 54; Indels 7; Gaps 7;

Db 131 PGNQGGPPQG-GPOQRP-OPNGQGGPPPGQPPQPPQGGNQGPP-OGGPH-PPR 186
:: |||:: |||:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ||
QY 202 AATQARPPHSGPRRLGCRANWHSYREAGVPLGAPARRGGGSASRLPLPKRPR 261
Db 187 PGNQGGPPPGQPPQPPQPPQGG-PPQAPRPGNQQGPPPGQPPQPPRTGNOQ 245
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 262 RGAPEPTPVQGSWAHPGTRGSPDRGFCVSPARPA-EAATSLGALSGTRSHPS 320
Db 246 -GPPQGGQPPRP 259
QY 321 VGRQHHAGPSTSRP 335

RESULT 11
ENTRY S11790 #type complete
TITLE nodulation protein nodD3 - Rhizobium leguminosarum bv.
phaseoli
ORGANISM #formal_name Rhizobium leguminosarum bv. phaseoli
DATE 21-Nov-1993 #sequence_revision 13-Jan-1995 #text_change
16-Feb-1997
ACCESSIONS S11790
REFERENCE S11786
#authors Davis, E.O.; Johnston, A.W.B.
#journal Mol. Microbiol. (1990) 4:921-932
#title Analysis of three nodD genes in Rhizobium leguminosarum
biovar phaseoli; nodD1 is preceded by nleE, a gene whose
product is secreted from the cytoplasm.
#cross-references MUID:91014692
#accession S11790
#status not compared with conceptual translation
#molecule_type DNA
#residues 1-302 #label DAV
GENETICS
#gene nodD3
CLASSIFICATION #superfamily regulatory protein lysR
KEYWORDS DNA binding; transcription regulation
SUMMARY #length 302 #molecular-weight 34274 #checksum 3819

Query Match 1.5%; Score 126; DB 2; Length 302;
Best Local Similarity 31.4%; Pred. No. 2.92e-02;
Matches 38; Conservative 28; Mismatches 46; Indels 9; Gaps 9;

Db 17 LMIEENLTAAFSINLSPAMSAAYRRL-RSYRDEL-FTMRGREFVTP-R-AEDLAPA 72
QY 9 LURTWEALAPATP-AMPAPRCRAVRSLRSHYREVLPLATEFVRLGPGQVRLVQGDPA 67
Db 73 IREALQHLRLNIPDKTTPQSDRHRF-VS-LCDFVTWVLFQKILRLAREAPGISFDL 130
QY 68 AFRALVQACLVCPWDA-RPPPAAPSFQVSCLEKELVARVL-QRLCERCAKNVLAFGAL 125
Db 131 L 131

```

```

QY 126 L 126

RESULT 12
ENTRY S10889 #type complete
TITLE proline-rich protein - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 07-Oct-1994 #sequence_revision 26-May-1995 #text_change
08-Sep-1997
ACCESSIONS S10889
REFERENCE S02127
#authors Lyons, K.M.; Stein, J.H.; Smithies, O.
#journal Genetics (1988) 120:267-278
#title Length polymorphisms in human proline-rich protein genes
generated by intragenic unequal crossing over.
#cross-references MUID:89121440
#accession S10889
#status Preliminary; translation not shown
#molecule_type DNA
#residues 1-309 #label LYO
#cross-references EMBL:X07881; NID:g35637; PID:g296669
GENETICS
#introns 22/1: 34/1
CLASSIFICATION #superfamily proline-rich protein
SUMMARY #length 309 #molecular-weight 30936 #checksum 3043

Query Match 1.5%; Score 129; DB 2; Length 309;
Best Local Similarity 30.4%; Pred. No. 1.24e-02;
Matches 45; Conservative 38; Mismatches 53; Indels 12; Gaps 12;

Db 122 GPPP-QGGNQQG-PPHPKPGEGPPP-QGG-NOSQGGPPRP-GKPE-GPPPQGGNQQG 175
||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 194 GPLYQLGAATQARPPHSGPRRLGCRANWHSYREAGVPLGAPARRGGGSASRS 253
Db 176 -P-PRCKPGEPPQGGNQQGPPRPKPGEGPPPQGGNQQGPPRPKPGEGSPSGG 233
QY 254 LPLKPRPRGAAPERTPVQGSWAHPGTRGSPDRG-F-CVVSPARPAB-EATSLGA 310
Db 234 -NKPRGPPPHKPGKQPPPGQGNKPPQRP 260
QY 311 LSGTRHSHPSVGRQHHAGPPTSRPPR 338

RESULT 13
ENTRY A28996 #type complete
TITLE proline-rich protein M14 precursor - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change
20-Mar-1998
ACCESSIONS A28996
REFERENCE A28996
#authors Ann, D.K.; Smith, K.; Carlson, D.M.
#journal J. Biol. Chem. (1988) 263:10887-10893
#title Molecular evolution of the mouse proline-rich protein
multigene family. Insertion of a long interspersed repeated
DNA element.
#cross-references MUID:88273214
#accession A28996
#molecule_type DNA
#residues 1-317 #label ANN
#cross-references GB:M23236; GB:J03891; NID:g200535; PID:g567232
GENETICS
#introns 22/1
CLASSIFICATION #superfamily proline-rich protein
KEYWORDS saliva
FEATURE
1-15 #domain signal sequence #status predicted #label SIG\
16-317 #product proline-rich protein M14 #status predicted
#label MAT
SUMMARY #length 317 #molecular-weight 31719 #checksum 8454

Query Match 1.5%; Score 132; DB 2; Length 317;

```

```
Best Local Similarity 26.4%; Pred. No. 5.17e-03;
Matches 39; Conservative 38; Mismatches 68; Indels 7; Gaps 6;

Db 37 SGSQPRPPVNGSQGPPPGGQPRP-PQGPPPPGQPRPPGQPPGQPRP-PQGP 94
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 202 AATQARPPPHAS--GPRRLGCE-RAWNHSVREAGVPLGLPAPGARRRGGSSASRSLPLPK 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 95 PPGGQPRPPGQPPPGG-QPRPPGQPPPGGQPRPPGQPPPGGQPRPPGQPPPP 153
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 259 RPRGAPEPRTVPVQGSWAHPGRTGRGSDRGFCVVSFAR-PAEATSLGALSCTRHS 317
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 154 GGPQPRPPGQPPPPAGQPRPPGQPPPP 181
QY 318 HPSVGRQHAGPSTSRPPRMDTPCPP 345

RESULT 14
ENTRY S22933 #type complete
TITLE testis-specific protein Bsl3 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
10-Sep-1997
ACCESSIONS S22933
REFERENCE S22933
#authors Mazarakis, N.D.; Nelki, D.; Lyon, M.F.; Ruddy, S.; Evans,
E.P.; Freemont, P.; Dudley, K.
#journal Development (1991) 111:561-571
#title Isolation and characterisation of a testis-expressed
developmentally regulated gene from the distal inversion of
the mouse t-complex.
#cross-references MUID:91372153
#accession S22933
#molecule_type mRNA
#residues 1-566 #label MAZ
#cross-references EMBL:X52128; NID:g54852; PID:g54853
SUMMARY #length 566 #molecular-weight 61970 #checksum 3815

Query Match 1.5%; Score 127; DB 2; Length 566;
Best Local Similarity 35.7%; Pred. No. 2.19e-02;
Matches 30; Conservative 19; Mismatches 30; Indels 5; Gaps 5;

Db 34 SARGTRVGSTVARPPSPGQPRGCAVKTAPRGVGHGLRTGPTSRCPQSARAKLPS 93
QY 239 PAPGARRRGGSSASRSLPL-PKPRRGAPEPRTVPVQGS-WAHP-GRTRGSDRG-FCV 294
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 94 VTRGAPLPPSPG-KHGLGTPSSH 116
QY 295 VSPARAEATSLGALSCTRHS 318
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
ENTRY S27923 #type complete
TITLE gene LF3 protein - human herpesvirus 4
ORGANISM #formal_name human herpesvirus 4, Epstein-Barr virus
DATE 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change
08-Sep-1997
ACCESSIONS S27923
REFERENCE S27923
#authors Parker, B.D.; Bankier, A.; Satchwell, S.; Barrell, B.;
Farrell, P.J.
#submission submitted to the EMBL Data Library, August 1990
#description Sequence and transcription of Raji Epstein-Barr virus DNA
spanning the B95-8 deletion region.
#accession S27923
#status preliminary
#molecule_type DNA
#residues 1-924 #label PAR
#cross-references EMBL:M35547; NID:g330420; PID:g330421
SUMMARY #length 924 #molecular-weight 94304 #checksum 8997

Query Match 1.5%; Score 132; DB 2; Length 924;
Best Local Similarity 28.2%; Pred. No. 5.17e-03;
Matches 37; Conservative 33; Mismatches 52; Indels 9; Gaps 9;
```

```
Db 38 AAPRAGPEPRTRLQPATPRRSGAADPADPVGHFAA-PRAPGPEPRTRLQ-PATPRRSGA 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 206 ARPPPHASGPRRLGCERAWNHSVREAGVPLGLPAPGARRRGGSSASRSLPLPKRPRGAA 265
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 96 ADPA-DPVGHFA-A-P-RAPGPEPRTRLQPATPRRSGAADPADPVGHFAA-PRAPGPEPRT 151
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 266 PEPRTVPVQGSWAHPGRTGRGSDRG-FCVVSFARP-AEEATSLGALSCTRHSHP-SVG 322
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 152 RLQPATPRRSG 162
QY 323 ROHHAGPESTS 333
```

Search completed: Thu Dec 24 07:48:05 1998
Job time : 320 secs.

M E S R E H
(TM)

Release 3.1A John F. Collins, Blocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 24 07:48:22 1998; MasPar time 38.16 Seconds
811.688 Million cell updates/sec

Tabular output not generated.

Title: >US-08-951-733-20
Description: (1-1154) from US08951733.pep
Perfect Score: 8624
Sequence: 1 HASGQRCVLLRTWEALPAT.....TALEAAANPALPSDFKTLTD 1154

Scoring table: PAM 150
Gap 11

Searched: 74019 seqs, 26840295 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swissprot

Statistics: Mean 55.898; Variance 96.260; scale 0.581

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|--------------|-------------------------|-----------|
| 1 | 162 | 1.9 | 660 | 1 YHL1_EBV | HYPOTHETICAL BHLF1 PRO | 4.64e-09 |
| 2 | 153 | 1.8 | 504 | 1 ATIN_HSVBP | ALPHA TRANS-INDUCING P | 1.47e-07 |
| 3 | 150 | 1.7 | 234 | 1 PRPM_HUMAN | SALIVARY PROLINE-RICH | 4.56e-07 |
| 4 | 142 | 1.6 | 276 | 1 PRPL_HUMAN | SALIVARY PROLINE-RICH | 8.72e-06 |
| 5 | 137 | 1.6 | 628 | 1 V70K_TYMV | 69 KD PROTEIN. | 5.26e-05 |
| 6 | 133 | 1.5 | 247 | 1 PRP4_HUMAN | SALIVARY PROLINE-RICH | 2.15e-04 |
| 7 | 131 | 1.5 | 296 | 1 PRP3_MOUSE | PROLINE-RICH PROTEIN D | 4.31e-04 |
| 8 | 126 | 1.5 | 302 | 1 NOD3_RHLPL | MODULATION PROTEIN D I | 2.37e-03 |
| 9 | 127 | 1.5 | 566 | 1 TS13_MOUSE | TESTIS-SPECIFIC PROTEIN | 1.69e-03 |
| 10 | 125 | 1.4 | 261 | 1 PRP2_MOUSE | PROLINE-RICH PROTEIN M | 3.32e-03 |
| 11 | 121 | 1.4 | 322 | 1 NOD1_RHLSN | MODULATION PROTEIN D I | 1.25e-02 |
| 12 | 118 | 1.4 | 399 | 1 BM8A_MOUSE | BONE MORPHOGENETIC PRO | 3.30e-02 |
| 13 | 125 | 1.4 | 402 | 1 BMP8_HUMAN | BONE MORPHOGENETIC PRO | 3.32e-02 |
| 14 | 118 | 1.4 | 431 | 1 UL61_HCMVA | HYPOTHETICAL PROTEIN U | 3.30e-02 |
| 15 | 125 | 1.4 | 501 | 1 MEHL_HUMAN | MESENCHYME FORK HEAD P | 3.32e-03 |
| 16 | 118 | 1.4 | 514 | 1 VE2_HPV05 | REGULATORY PROTEIN E2. | 3.30e-02 |
| 17 | 114 | 1.3 | 202 | 1 VG01_HSVB | HYPOTHETICAL GENE 1 PR | 1.17e-01 |
| 18 | 116 | 1.3 | 270 | 1 MIAE_SALVE | TRNA-(MS[2]IO[6]A)-HYD | 6.23e-02 |
| 19 | 115 | 1.3 | 279 | 1 Y091_NPVOP | HYPOTHETICAL 29.3 KD P | 8.54e-02 |
| 20 | 109 | 1.3 | 301 | 1 NDO3_RHLO | MODULATION PROTEIN D3. | 5.40e-01 |
| 21 | 114 | 1.3 | 321 | 1 NOD1_BRAJA | MODULATION PROTEIN D I | 1.17e-01 |
| 22 | 109 | 1.3 | 324 | 1 MATD_NEUCR | MATING TYPE PROTEIN A- | 5.40e-01 |
| 23 | 115 | 1.3 | 331 | 1 PRP1_HUMAN | SALIVARY PROLINE-RICH | 8.54e-02 |

| | | | | | | |
|----|-----|-----|------|---------------|------------------------|----------|
| 24 | 114 | 1.3 | 408 | 1 CGEL_XENLA | GL/S-SPECIFIC CYCLIN E | 1.17e-01 |
| 25 | 114 | 1.3 | 408 | 1 CGE2_XENLA | GL/S-SPECIFIC CYCLIN E | 1.17e-01 |
| 26 | 110 | 1.3 | 408 | 1 CGE3_XENLA | GL/S-SPECIFIC CYCLIN E | 3.99e-01 |
| 27 | 111 | 1.3 | 498 | 1 VE2_HPV08 | REGULATORY PROTEIN E2. | 2.95e-01 |
| 28 | 111 | 1.3 | 509 | 1 VE2_HPV36 | REGULATORY PROTEIN E2. | 2.95e-01 |
| 29 | 113 | 1.3 | 514 | 1 VE2_HPV5B | REGULATORY PROTEIN E2. | 1.59e-01 |
| 30 | 110 | 1.3 | 628 | 1 V70K_TYMV | 69 KD PROTEIN. | 3.99e-01 |
| 31 | 112 | 1.3 | 633 | 1 LAL17_YEAST | PROLINE-RICH PROTEIN L | 2.17e-01 |
| 32 | 113 | 1.3 | 704 | 1 SYN1_RAT | SYNAPSINS IA AND IB. | 1.59e-01 |
| 33 | 114 | 1.3 | 705 | 1 SYN1_HUMAN | SYNAPSINS IA AND IB (B | 1.17e-01 |
| 34 | 108 | 1.3 | 706 | 1 SYN1_BOVIN | SYNAPSINS IA AND IB. | 7.28e-01 |
| 35 | 113 | 1.3 | 1255 | 1 MUC1_HUMAN | MUCIN 1 PRECURSOR (POL | 1.59e-01 |
| 36 | 115 | 1.3 | 3149 | 1 TEGU_EBV | LARGE TEGUMENT PROTEIN | 8.54e-02 |
| 37 | 107 | 1.2 | 174 | 1 BARL_CHITE | BALBIANI RING PROTEIN | 9.79e-01 |
| 38 | 106 | 1.2 | 233 | 1 YFUR_ECOLI | HYPOTHETICAL TRANSCRIP | 1.31e+00 |
| 39 | 106 | 1.2 | 285 | 1 YAEY_ECOLI | HYPOTHETICAL TRANSCRIP | 1.31e+00 |
| 40 | 107 | 1.2 | 318 | 1 NODD_RHLT | MODULATION PROTEIN D. | 9.79e-01 |
| 41 | 107 | 1.2 | 494 | 1 MEHL_MOUSE | MESENCHYME FORK HEAD P | 9.79e-01 |
| 42 | 107 | 1.2 | 529 | 1 DNE2_ADE05 | EARLY E2A DNA-BINDING | 9.79e-01 |
| 43 | 105 | 1.2 | 863 | 1 HIS2_NEUCR | PHOSPHORIBOSYL-AMP CYC | 1.76e+00 |
| 44 | 105 | 1.2 | 890 | 1 ACO1_ECOLI | ACONITATE HYDRATASE 1 | 1.76e+00 |
| 45 | 105 | 1.2 | 2774 | 1 MAPA_RAT | MICROTUBULE-ASSOCIATED | 1.76e+00 |

ALIGNMENTS

| RESULT | 1 | STANDARD; | PRT; | 660 AA. |
|--------|---|-----------|------|---------|
| ID | YHL1_EBV | | | |
| AC | P03181 | | | |
| DT | 21-JUL-1986 (REL. 01, CREATED) | | | |
| DT | 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE) | | | |
| DT | 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE) | | | |
| DE | HYPOTHETICAL BHLF1 PROTEIN. | | | |
| OS | EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4). | | | |
| OC | VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; GAMMAHERPESVIRINAE. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE; 84270667. | | | |
| RA | BAER R., BANKIER A.T., BIGGIN M.D., DEININGER P.L., FARRELL P.J., GIBSON T.J., HATFULL G., HUDSON G.S., SATCHWELL S.C., SEGUIN C., TUFFNELL P.S., BARRELL B.G.; | | | |
| RL | NATURE 310:207-211(1984). | | | |
| DR | EMBL; V01555; -; NOT_ANNOTATED_CDS. | | | |
| DR | PIR; A03742; Q0BE3. | | | |
| KW | HYPOTHETICAL PROTEIN; EARLY PROTEIN; REPEAT. | | | |
| FT | DOMAIN 149 648 4 X 125 AA TANDEM REPEATS. | | | |
| FT | REPEAT 149 273 1. | | | |
| FT | REPEAT 274 398 2. | | | |
| FT | REPEAT 399 523 3. | | | |
| FT | REPEAT 524 648 4. | | | |
| SQ | SEQUENCE 660 AA: 66244 MW; 372F08C5 CRC32; | | | |

| | | | | |
|-----------------------|--------------------------------|---|----------------|---------------------|
| Query Match | 1.9%; | Score 162; | DB 1; | Length 660; |
| Best Local Similarity | 30.8%; | Pred. No. 4.64e-09; | | |
| Matches | 48; | Conservative 34; | Mismatches 60; | Indels 14; Gaps 12; |
| Db | 251 | GGPPTRSAGAAQ-RTHRRPPGCPGRARNPGCPRTWRR--R-SGAQRGHPPPGAGQRPSP 306 | | |
| QY | 194 | GPPLYQLGRTAQTQARPPPHASG-PRR-R-LGCRANNNHSREAGVPLGLPAPGA-RRRGGS 249 | | |
| Db | 307 | TGGRPAAPGAPGTTPAAGPGGGAAVSPGATPHPRGSGPADPP--AAARLPPPERQEPRLP 364 | | |
| QY | 250 | ASRSLPLPKRRRGAPEPER-TPVGQGSWAHPGRTGRPGSDRGFCVVSPARPAEATSL 308 | | |
| Db | 365 | QDLAAQRC-PAGPPPTRSQA-AAQFTHR-RPPGCP 397 | | |
| QY | 309 | GALSGTRSHSPSVGRQHHAGPSTSPRPWDTPCP 344 | | |
| RESULT | 2 | STANDARD; | PRT; | 504 AA. |
| ID | ATIN_HSVBP | | | |
| AC | P30020; | | | |
| DT | 01-APR-1993 (REL. 25, CREATED) | | | |

DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE ALPHA TRANS-INDUCING PROTEIN (ALPHA-TIF).
 OS BOVINE HERPESVIRUS TYPE 1 (STRAIN P8-2).
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93012995.
 RA CARPENTER D.E., MISRA V.;
 RL GENE 119:259-263(1992).
 CC -1- FUNCTION: RESPONSIBLE FOR TRANSCRIPTIONAL ACTIVATION OF IMMEDIATE
 CC EARLY PROMOTERS (ALPHA GENES) (BY SIMILARITY).
 CC -1- SIMILARITY: TO OTHER HERPESVIRUSES ALPHA TRANS-INDUCING PROTEIN.
 DR EMBL; Z11610; E264419; -.
 DR PIR; S24229; S24229.
 DR PIR; JCI1306; JCI1306.
 KW TRANSCRIPTION REGULATION; TRANS-ACTING FACTOR; DNA-BINDING.
 SQ SEQUENCE 504 AA; 54028 MW; 79F42020 CRC32;
 Query Match 1.8%; Score 153; DB 1; Length 504;
 Best Local Similarity 31.3%; Pred. No. 1.47e-07;
 Matches 46; Conservative 33; Mismatches 56; Indels 12; Gaps 11;

Db 351 APAEAGGWRSGSTRTRGAAARSTTGLQPCGPRRAKC-CRATP-RQLR--ARGE 406
 QY 208 PPHASGPRRLGGERAWNHSVREAGVPLGLPAPGARRGGSASRSLPKRPRGAAPE 267
 Db 407 PRHTS-GSGAFSQ-GRPRGVRCLGWACKARSGPARGPGSPVRSGLGLSR-ARGSPGP 463
 QY 268 PERTPVGGGWAHPGTRGTPSDR-GF-CVV-S-PARAEAEATSGALSGTRHSPSVGR 323
 Db 464 GPACGFSRAGGRRRASPAMP-FGGT 489
 QY 324 QHHAGPSTSRPPRPWDTCPVPVYAE 350

RESULT 3
 ID PRPM_HUMAN STANDARD; PRT; 234 AA.
 AC P10161; P02813;
 DT 01-MAR-1989 (REL. 10, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE SALIVARY PROLINE-RICH PROTEIN PO (ALLELE M) (CONTAINS: PEPTIDE P-D)
 DE (FRAGMENT).
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 89121439.
 RA LYONS K.M., STEIN J.H., SMITHIES O.;
 RL GENETICS 120:255-265(1988).
 RN [2]
 RP SEQUENCE OF 165-234.
 RX MEDLINE; 83186122.
 RA SAITOH E., ISEMURA S., SANADA K.;
 RL J. BIOCHEM. 93:495-502(1983).
 DR EMBL; X07704; E265547; -.
 DR PIR; A03295; PIHUSD.
 DR PIR; S03175; S03175.
 DR HSSP; P19999; ICLG.
 DR MIM; 168730; -.
 DR MIM; 180990; -.
 KW REPEAT; PAROTID GLAND; SALIVA; MULTIGENE FAMILY.
 FT NON_TER 1 1
 FT CHAIN 165 234 PEPTIDE P-D.
 SQ SEQUENCE 234 AA; 23676 MW; ED2B4ADC CRC32;

Query Match 1.7%; Score 150; DB 1; Length 234;
 Best Local Similarity 29.3%; Pred. No. 4.56e-07;
 Matches 43; Conservative 41; Mismatches 53; Indels 10; Gaps 10;

Db 88 GPPP-OGGNQSQG-PPPHGKPERPPP-OGG-NQSHRPPPPP-GKPER-PPPOGGNOSQG 141
 QY 194 GPPLYQLGAATQARPPPHASGPRRLGGERAWNHSVREAGVPLGLPAPGARRRGGASRS 253
 Db 142 -P-PPHPGPEGPQEGNKSARSPPGKPGQPGQEGNKPQGGPPPGKPGQPGPPPGGN 199
 QY 254 LPLPKRPRGAAPERTPVGGGWAHPGTRGTPSDR-GFCVVSPPARPAE-EATSGAL 311
 Db 200 PQQPQAPAGKPGQPPPPPGGRRPRP 226
 QY 312 SGRHSHPSVGRQHHAGPPTSRPRP 338
 RESULT 4
 ID PRPL_HUMAN STANDARD; PRT; 276 AA.
 AC P10162; P02813;
 DT 01-MAR-1989 (REL. 10, CREATED)
 DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
 DE SALIVARY PROLINE-RICH PROTEIN PO (ALLELE K) (CONTAINS: PEPTIDE P-D)
 DE (FRAGMENT).
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 89121439.
 RA LYONS K.M., STEIN J.H., SMITHIES O.;
 RL GENETICS 120:255-265(1988).
 RN [2]
 RP SEQUENCE OF 207-276.
 RX MEDLINE; 83186122.
 RA SAITOH E., ISEMURA S., SANADA K.;
 RL J. BIOCHEM. 93:495-502(1983).
 DR EMBL; X07715; E4806; ALT_SEQ.
 DR PIR; A03295; PIHUSD.
 DR PIR; S03176; S03176.
 DR HSSP; P19999; ICLG.
 DR MIM; 168730; -.
 DR MIM; 180990; -.
 KW REPEAT; PAROTID GLAND; SALIVA; MULTIGENE FAMILY.
 FT NON_TER 1 1
 FT CHAIN 207 276 PEPTIDE P-D.
 SQ SEQUENCE 276 AA; 27816 MW; 4838945A CRC32;
 Query Match 1.6%; Score 142; DB 1; Length 276;
 Best Local Similarity 28.2%; Pred. No. 8.72e-06;
 Matches 40; Conservative 41; Mismatches 52; Indels 9; Gaps 9;

Db 134 QGGNOSQG-PPPHGKPERPPP-OGG-NQSHRPPPPP-GKPER-PPPOGGNOSQG-P-PP 186
 QY 199 QLGAATQARPPPHASGPRRLGGERAWNHSVREAGVPLGLPAPGARRRGGASRSPLPK 258
 Db 187 HPKPEGPPPEGNKSARSPPGKPGQPGQEGNKPQGGPPPGKPGQPGPPPGGNPQQPO 246
 QY 259 RPRGAAPERTPVGGGWAHPGTRGTPSDR-GFCVVSPPARPAE-EATSGALSGTRH 316
 Db 247 APGAKPGQPPPPPGGRRPRP 268
 QY 317 SHPSVGRQHHAGPPTSRPRP 338
 RESULT 5
 ID V70K_TYMV STANDARD; PRT; 628 AA.
 AC P10357;
 DT 01-MAR-1989 (REL. 10, CREATED)
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
 DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
 DE 69 KD PROTEIN.
 OS TURNIP YELLOW MOSAIC VIRUS.
 OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; TYMOVIRIDAE.
 RN [1]

```

SEQUENCE FROM N.A.
RX MEDLINE; 88289359.
RA MORCH M.D., BOYER J.C., HAENNI A.L.;
RL NUCLEIC ACIDS RES. 16:6157-6173(1988).
CC -I- FUNCTION: NOT KNOWN.
CR -I- SIMILARITY: TO 65 TO 70 KD PROTEIN FROM OTHER TMOVIRUSES.
DR ENBL; X07441; G62223; ALT_SEQ.
DR PIR; S01955; S01955.
SQ SEQUENCE 628 AA; 69195 MW; 9E64ED49 CRC32;

Query Match      1.6%; Score 137; DB 1; Length 628;
Best Local Similarity 25.0%; pred. No. 5.26e-05;
Matches 73; Conservative 64; Mismatches 134; Indels 21; Gaps 20;

Db 157 GPLYETKPTSVQPSATRGPSFRPILLPKV-VH-VHDDPHSSLSRPGSSRSRQLPT 214
Qy 194 GPDYQLGAATQARPPPHAS-GPR-RKLGGERANNWSVREAGVPLGLPAPGAR-RR-GGS 249
Db 215 VRRPLLAPNOFHSRPPPLSDPDGILGPRPLAPHSTRDPPRPPI-TPGGSN-THDLRPL 272
Qy 250 ASRSLPLKPRRGAAPERTPVG-QGSA-HPGRTGPDGRGFCVSPARPAEATSL 307
Db 273 SVLPRTSPRGLLPNRRHTSTGHPPTTSRTGPPSRQLRPVH-LYQSSPHTNFRP 331
Qy 308 EG-ALSTRTSHSPVGRHAGPPSTSRPRPRMDTCPVYAEIKHFLY-SSGDKQLRP 365
Db 332 SSIRKDALLOTGPRGLHLELGPANLRTSGSPPTKRRPLRSESPNRLPKPPEATLAP 391
Qy 366 SFLL-SSL-R-PSLTGARRLVETIFLGSRPWPMPGTPRRLPRLPQRYQWRPFLLELLG- 420
Db 392 SYRHRRYPYLLNPDAALPSIATYSSRGKTHHSLPKGALPK-EGAPPPPPRL 442
Qy 421 NHAQC-EYGVLLKTHCPFLRA-AVTPAGVCAREKPGQSVAAPEEDTPRRL 470

```

[illegible]

Query Match 1.5%; Score 133; DB 1; Length 247;
Best Local Similarity 27.4%; Pred. No. 2.15e-04;
Matches 40; Conservative 40; Mismatches 56; Indels 10; Gaps 10;

[illegible]

| | | | | | | |
|----|-----------------------|--|---------------------|-------|----------------|-------------------|
| | Query Match | 1.5% | Score 131; | DB 1; | Length 296; | |
| | Best Local Similarity | 26.4%; | Pred. No. 4.31e-04; | | | |
| | Matches | 39; | Conservative | 38; | Mismatches 64; | Indels 7; Gaps 6; |
| Db | 16 | SGSQPRPPVNGSQGPPPPCGQPQRP-PQGPPPPGGQPRPPOGPPPPGPGQPPRP-PQGP 73 | | | | |
| QY | 202 | AATQARPPPHAS--GPRRRLLGCE-RAWNHSVRAGVPLGFAPGARRRGGSARSRLPLPK 258 | : | : | : | : |
| Db | 74 | PPGCGPQRPQPPGPPPPGG-QPQRPPQGPQPPGCGQPQRPPOGPPPPGPGQPPRPQGPpppp 132 | : | : | : | : |
| QY | 259 | RPRGAAPERTEPVGGSWAHFGRTRGSDRGFCVVSPAR-PAEATSLEGALSGTRHS 317 | : | : | : | : |
| Db | 133 | GGQPQRPQGPppppPAGQPQRPQGPpppp 160 | : | : | : | : |
| QY | 318 | HPSVGROHHAGPPSTSRPPRPWDTPCPP 345 | : | : | : | : |

RESULT 8
ID NOD3_RHILP STANDARD; PRRT; 302 AA.
AC P23720;
DT 01-NOV-1991 (REL. 20, CREATED)
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE NODULATION PROTEIN D III.
GN NODD3.
OS RHIZOBIUM LEGUMINOSARUM (BIOVAR PHASEOLI).
OG PLASMID SYM.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC RHIZOBIACEAE.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-8002;
RC MEDLINE; 91014592.
RX DAVIS E.O.; JOHNSTON A.W.B.;
RA MOL. MICROBIOL. 4:921-932(1990).
RL -1- FUNCTION: NODD REGULATES THE EXPRESSION OF THE NOD ABCFE GENES
CC WHICH ENCODE OTHER NODULATION PROTEINS. NODD IS ALSO A NEGATIVE
CC REGULATOR OF ITS OWN EXPRESSION. BINDS FLAVONOIDS AS INDUCERS.
CC

Dd 94 VTRGAPLPSPG-KHGGTGTPSSH 116
| : | : | : | : | : ||
Qy 295 VSPARPAEEATSLLEGALSGTRHSH 318

RESULT 10
ID PRP2_MOUSE STANDARD; PRT; 261 AA.
AC P05142;
Dt 13-AUG-1987 (REL. 05, CREATED)
Dt 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
Dt 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE PROLINE-RICH PROTEIN MP-2 PRECURSOR.
GN PRP.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86059475.
Rr ANN D.K., CARLSON D.M.;
RL J. BIOL. CHEM. 260:15863-15872(1985).
DR ENBL; M12099; G200547; -
DR HSSP; P19999; ICLG.
KW REPEAT; SALIVA; SIGNAL.
FT SIGNAL 1 15
FT CHAIN 16 261
SQ SEQUENCE 261 AA; 26034 MW; 9D830DAF CRC32;

Query Match 1.4%; Score 125; DB 1; Length 261;
Best Local Similarity 27.0%; Pred. No. 3.32e-03;
Matches 40; Conservative 33; Mismatches 69; Indels 6; Gaps

Dd 76 PPGPPPPGGQPQRPPQGPPPGQPQRPPGPQQRRPPGPQPGRPPGPPGPPGPPGPPGPP 1
Qy 203 ATGARPPPHASGRPRLRGERAWNHSVRAGVPLGLDPAGARRGGSASRSALUPLPKRRPR 2

Dd 136 PPPPGGQLRPPGPPPPPAGP-QPRPPGPPFPAGPQPRPPGPPPTTGPQRPRTQGPPPT 1
Qy 263 GAAPPE-PE-RTVPQGSWAHPGKRGPSDRGCVCVSPARPAAEATSLEGALSGRHRSHPS 3

Dd 195 GGQQQQPPGPPPPGPPGPPPPGPPPP 222
Qy 321 VGRQHHA--GPPSTSRP-PRPWDTCPP 345

RESULT 11
ID NOD1_RHISN STANDARD; PRT; 322 AA.
AC P53359;
Dt 01-NOV-1997 (REL. 35, CREATED)
Dt 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
Dt 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MODULATION PROTEIN D I.
GN NODDI OR Y4L.
OS RHIZOBIUM SP. (STRAIN NGR234).
OG PLASMID SYM PNGR234A.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC RHIZOBIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97305956.
Rr FREIBERG C.A., FELLAY R., BAIROCH A., BROUGHTON W.J., ROSENTHAL A., PERRET X.;
RL NATURE 387:1394-401(1997).
CC -I- FUNCTION: NODD REGULATES THE EXPRESSION OF THE NOD ABCFE GENES WHICH ENCODE OTHER MODULATION PROTEINS. NODD IS ALSO A NEGATIVE REGULATOR OF ITS OWN EXPRESSION. BINDS FLAVONOIDS AS INDUCERS.
CC -I- SIMILARITY: BELONGS TO THE LYSE FAMILY OF TRANSCRIPTIONAL REGULATORS.

Dd ENBL; A5000065; G2182301; -
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
RW MODULATION; TRANSCRIPTION REGULATION; DNA-BINDING; ACTIVATOR;
KW REPRESSOR; MULTIGENE FAMILY; PLASMID.
FW DNA_BIND 23 42 H-T-H MOTIF (BY SIMILARITY).

| | |
|-----------------------|---|
| AC | F34820; |
| AD | 01-FEB-1994 (REL. 28, CREATED) |
| DT | 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE) |
| DT | 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE) |
| DE | BONE MORPHOGENETIC PROTEIN 8 PRECURSOR (BMP-8) (OSTEOGENIC PROTEIN 2) |
| DE | (OP-2). |
| GN | BMP8. |
| OS | HOMO SAPIENS (HUMAN). |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; |
| QC | EUTHERIA; PRIMATES. |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | TISSUE=HIPPOCAMPUS; |
| RX | MEDLINE; 93094231. |
| RA | OZKAYNAK E., SCHNEGELSBERG P.N.J., JIN D.F., CLIFFORD G.M., |
| RA | WARREN F.D., DRIER E.A., OPPERMAN H.; |
| JL | J. BIOL. CHEM. 267:25220-25227(1992). |
| CC | -I- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION. MAY BE THE |
| CC | OSTEOINDUCTIVE FACTOR RESPONSIBLE FOR THE PHENOMENON OF |
| CC | EPITHELIAL OSTEOGENESIS. PLAYS A ROLE IN CALCIUM REGULATION |
| CC | AND BONE HOMEOSTASIS (BY SIMILARITY). |
| CC | -I- SUBUNIT: HOMODIMER, DISULFIDE-LINKED. |
| CC | -I- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. |
| DR | EMBL; M97016; GI89390; - |
| DR | .PIR; A45056; A45056. |
| DR | HSSP; P08112; 1TFG. |
| DR | MIM; 602284; - |
| KW | PROSITE; PS00250; TGF_BETA; 1. |
| FT | SIGNAL; GROWTH FACTOR; CYTOKINE; BONE; CARTILAGE; GLYCOPROTEIN. |
| FT | SIGNAL 1 19 POTENTIAL. |
| FT | PROPEP 20 263 POTENTIAL. |
| FT | CHAIN 264 402 BONE MORPHOGENETIC PROTEIN 8. |
| FT | DISULFID 301 367 BY SIMILARITY. |
| FT | DISULFID 330 399 BY SIMILARITY. |
| FT | DISULFID 334 401 BY SIMILARITY. |
| FT | DISULFID 366 366 INTERCHAIN (BY SIMILARITY). |
| FT | CARBOHYD 158 158 POTENTIAL. |
| FT | CARBOHYD 343 343 POTENTIAL. |
| SQ | SEQUENCE 402 AA; 44764 MW; 8F20C81B CRC32; |
| Query Match | 1.4%; Score 125; DB 1; Length 402; |
| Best Local Similarity | 39.7%; Pred.No.3.32e-03; |
| Matches | 25; Conservative 12; Mismatches 21; Indels 5; Gaps |
| Db | 12 GLALCALGGSGPLRPP-GC-PQRRLG-ARERRDVQRETLAVLGPGR-PRPRAPPAAS 67 |
| QY | 194 GPLYLQGAATQA-RPPHASGPRRLGGERAWNSVREAGVPLGLPAGARRRGGSASR 152 : :::: : : : : : : |
| Db | 68 RLP 70 |
| QY | 253 SLP 255 |
| RESULT 14 | |
| ID | U161_HCMVA STANDARD; PRT; 431 AA. |
| AC | P16818; |
| DT | 01-AUG-1990 (REL. 15, CREATED) |
| DT | 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE) |
| DT | 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE) |
| DE | HYPOTHETICAL PROTEIN U161. |
| GN | U161. |
| OS | HUMAN CYTOMEGALOVIRUS (STRAIN AD169). |
| QC | VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE. |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE; 90269039. |
| RA | CHEE M.S., BANKIER A.T., BECK S., BOHNI R., BROWN C.M., CERNY R., |
| RA | HORSNELL T., HUTCHISON C.A. III, KOUZARIDES T., MARTIGNETTI J.A., |
| RA | FREDIE E., SATCHELL S.C., TOMLINSON P., WESTON K.W., BARRELL B.G.; |
| RL | CURR. TOP. MICROBIOL. IMMUNOL. 154:125-169(1990). |
| DR | EMBL; X17403; E27260; - |
| DR | .PIR; S09824; S09824. |
| KW | HYPOTHETICAL PROTEIN. |

[illegible]

M P E R E H
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 24 07:51:54 1998; MasPar time 76.57 Seconds
750.503 Million cell updates/sec

Tabular output not generated.

Title: >US-08-951-733-20
Description: (1-1154) from US08951733.pep
Perfect score: 8624
Sequence: 1 HASGQRCVLLRTWEALAPAT.....TALEAAANPALPSDFKTLTD 1154

Scoring table: PAM 150
Gap 11

Searched: 165420 seqs, 49795644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrembl6
1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human
5:sp.invertebrate 6:sp.mammal 7:sp.mhc 8:sp.organelle
9:sp.phage 10:sp.plant 11:sp.rodent 12:sp.unclassified
13:sp.vertibrate 14:sp.virus

Statistics: Mean 54.168; Variance 101.474; scale 0.534

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | | | | | | | |
|------------|-------|-------------|--------------|-------------------------------|-----------|--|--|
| Result No. | Score | Query Match | Length DB ID | Description | Pred. No. | | |
| 1 | 8465 | 98.2 | 1132 4 | O14746 TELOMERASE REVERSE TRA | 0.00e+00 | | |
| 2 | 8459 | 98.1 | 1132 4 | O14783 TELOMERASE CATALYTIC S | 0.00e+00 | | |
| 3 | 5161 | 59.8 | 1122 11 | O70372 TELOMERASE REVERSE TRA | 0.00e+00 | | |
| 4 | 678 | 7.9 | 988 3 | O13339 TELOMERASE REVERSE TRA | 1.37e-110 | | |
| 5 | 681 | 7.9 | 989 3 | O13338 TELOMERASE REVERSE TRA | 3.12e-111 | | |
| 6 | 439 | 5.1 | 1031 5 | O00939 TELOMERASE SUBUNIT P12 | 1.97e-60 | | |
| 7 | 424 | 4.9 | 67 11 | O35432 TELOMERASE CATALYTIC S | 2.25e-57 | | |
| 8 | 314 | 3.6 | 884 3 | O06163 CHROMOSOME XII COSMID | 1.71e-35 | | |
| 9 | 147 | 1.7 | 296 14 | O69118 HYPOTHETICAL PROTEIN (| 7.18e-06 | | |
| 10 | 143 | 1.7 | 574 3 | O36027 WISKOTT-ALDRICH SYNDRO | 2.85e-05 | | |
| 11 | 142 | 1.6 | 234 4 | O00600 PAROTID 'O' PROTEIN (F | 4.02e-05 | | |
| 12 | 141 | 1.6 | 276 2 | O04612 IS 1222 GENE ORF-A AND | 5.65e-05 | | |
| 13 | 136 | 1.6 | 316 4 | O99076 HOMEOBOX PROTEIN (FRAG | 3.03e-04 | | |
| 14 | 140 | 1.6 | 373 2 | O54226 POLYKETIDE SYNTHASE (F | 7.93e-05 | | |
| 15 | 135 | 1.6 | 585 14 | O41935 HYPOTHETICAL 60.2 KD P | 4.23e-04 | | |
| 16 | 136 | 1.6 | 897 11 | O70495 PLENTY-OF-PROLINES-101 | 3.03e-04 | | |
| 17 | 129 | 1.5 | 264 14 | O85301 HOMOLOGUE OF RETROVIRA | 3.01e-03 | | |
| 18 | 131 | 1.5 | 300 11 | O61888 PROLINE RICH PROTEIN. | 1.57e-03 | | |
| 19 | 129 | 1.5 | 309 4 | O04188 SALIVARY PROLINE-RICH | 3.01e-03 | | |
| 20 | 132 | 1.5 | 317 11 | O62103 PROLINE RICH PROTEIN P | 1.14e-03 | | |

| | | | | | | | |
|----|-----|-----|---------|-------------------------------|----------|--|--|
| 21 | 128 | 1.5 | 527 2 | O53913 BETA-KETOACYL SYNTHASE | 4.15e-03 | | |
| 22 | 133 | 1.5 | 539 5 | O02123 SIMILARITY TO COLLAGEN | 8.18e-04 | | |
| 23 | 132 | 1.5 | 924 14 | O99307 LF3 PROTEIN. | 1.14e-03 | | |
| 24 | 130 | 1.5 | 1236 2 | O06264 HYPOTHETICAL 131.9 KD | 2.18e-03 | | |
| 25 | 126 | 1.5 | 3247 14 | O65553 UL36. | 7.85e-03 | | |
| 26 | 117 | 1.4 | 197 5 | O17626 C04G2.8. | 1.27e-01 | | |
| 27 | 118 | 1.4 | 203 5 | O91497 CODED FOR BY C. ELEGAN | 9.42e-02 | | |
| 28 | 125 | 1.4 | 227 11 | O62107 PROLINE-RICH SALIVARY | 1.08e-02 | | |
| 29 | 118 | 1.4 | 236 14 | O85028 SUBSTRATE OF THE PROTE | 9.42e-02 | | |
| 30 | 118 | 1.4 | 322 2 | O83180 MPK3030 NODDI GENE IN | 9.42e-02 | | |
| 31 | 118 | 1.4 | 405 2 | O68033 EXONUCLEASE SBCD HOMOL | 9.42e-02 | | |
| 32 | 118 | 1.4 | 437 2 | O52256 FLHF. | 9.42e-02 | | |
| 33 | 123 | 1.4 | 439 10 | O42421 CHITINASE PRECURSOR. | 2.02e-02 | | |
| 34 | 124 | 1.4 | 464 10 | O41645 EXTENSIN (FRAGMENT). | 1.48e-02 | | |
| 35 | 123 | 1.4 | 494 4 | O15220 PRPL-2 PROTEIN. | 2.02e-02 | | |
| 36 | 125 | 1.4 | 503 4 | O43516 WASP INTERACTING PROTE | 1.08e-02 | | |
| 37 | 118 | 1.4 | 526 14 | O85027 VIRAL PROTEINASE. | 9.42e-02 | | |
| 38 | 122 | 1.4 | 628 5 | O22835 SP16 ISOLOG. | 2.76e-02 | | |
| 39 | 124 | 1.4 | 820 4 | O60585 SER/ARG-RELATED NUCLEA | 1.48e-02 | | |
| 40 | 115 | 1.3 | 106 14 | O41981 HYPOTHETICAL 9.9 KD PR | 2.32e-01 | | |
| 41 | 115 | 1.3 | 260 11 | O64306 PROLINE-RICH PROTEIN. | 1.72e-01 | | |
| 42 | 116 | 1.3 | 295 11 | O07611 PROLINE-RICH PROTEOGLY | 2.32e-01 | | |
| 43 | 115 | 1.3 | 297 4 | O16038 PRB1M PROTEIN (FRAGEN | 2.32e-01 | | |
| 44 | 115 | 1.3 | 382 4 | O00599 CON1. | 2.32e-01 | | |
| 45 | 116 | 1.3 | 488 10 | O04528 SEQUENCE OF BAC F20P5 | 1.72e-01 | | |

ALIGNMENTS

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|------|------|-----|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
| RESULT | 1 | | | | | | | | | | | | | | | | | | | |
| ID | O14746 | PRELIMINARY; | PRT; | 1132 | AA. | | | | | | | | | | | | | | | |
| AC | O14746; | | | | | | | | | | | | | | | | | | | |
| DT | 01-JAN-1998 | (TREMBLREL. 05, CREATED) | | | | | | | | | | | | | | | | | | |
| DT | 01-JAN-1998 | (TREMBLREL. 05, LAST SEQUENCE UPDATE) | | | | | | | | | | | | | | | | | | |
| DT | 01-JUN-1998 | (TREMBLREL. 06, LAST ANNOTATION UPDATE) | | | | | | | | | | | | | | | | | | |
| DE | TELOMERASE REVERSE TRANSCRIPTASE. | | | | | | | | | | | | | | | | | | | |
| GN | HRT. | | | | | | | | | | | | | | | | | | | |
| OS | HOMO SAPIENS (HUMAN). | | | | | | | | | | | | | | | | | | | |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; | | | | | | | | | | | | | | | | | | | |
| OC | EUTHERIA; PRIMATES. | | | | | | | | | | | | | | | | | | | |
| RN | [1] | | | | | | | | | | | | | | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | | | | | | | | | | | | | | |
| RC | TISSUE-KIDNEY. | | | | | | | | | | | | | | | | | | | |
| RX | MEDLINE; 97400623. | | | | | | | | | | | | | | | | | | | |
| RA | NAKAMURA T.M., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H., | | | | | | | | | | | | | | | | | | | |
| RA | LINGNER J., HARLEY C.B., CECI T.R.; | | | | | | | | | | | | | | | | | | | |
| RL | SCIENCE 277:955-959(1997). | | | | | | | | | | | | | | | | | | | |
| DR | EMBL; AF015950; G2330017. | | | | | | | | | | | | | | | | | | | |
| KW | RNA-DIRECTED DNA POLYMERASE. | | | | | | | | | | | | | | | | | | | |
| SQ | SEQUENCE 1132 AA; 126995 MW; 2DFBEDF3 CRC32; | | | | | | | | | | | | | | | | | | | |
| Query Match 98.2%; Score 8465; DB 4; Length 1132; | | | | | | | | | | | | | | | | | | | | |
| Best Local Similarity 100.0%; Pred. No. 0.00e+00; | | | | | | | | | | | | | | | | | | | | |
| Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | | | | | | | | | | | | |
| Db | 1 | MPRAPRCRAVRSLLRSHYREVLPATFVRLPGQWRLVQRGDPAAFRALVAQCLVCPW | 60 | | | | | | | | | | | | | | | | | |
| QY | 23 | MPRAPRCRAVRSLLRSHYREVLPATFVRLPGQWRLVQRGDPAAFRALVAQCLVCPW | 82 | | | | | | | | | | | | | | | | | |
| Db | 61 | DARPPPAASFQVSCLEKELVAVLQRLCERGAKNVLAFGFALLDARGGPPPAFTTSVR | 120 | | | | | | | | | | | | | | | | | |
| QY | 83 | DARPPPAASFQVSCLEKELVAVLQRLCERGAKNVLAFGFALLDARGGPPPAFTTSVR | 142 | | | | | | | | | | | | | | | | | |
| Db | 121 | SYLPNTVTDALRGSGAWGLLLRRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLVOLGA | 180 | | | | | | | | | | | | | | | | | |
| QY | 143 | SYLPNTVTDALRGSGAWGLLLRRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLVOLGA | 202 | | | | | | | | | | | | | | | | | |
| Db | 181 | ATQARPPPHASGPRRLGGERAWNHSVRAGVPLGLPAPGARRRGSGASRSLPKRPRR | 240 | | | | | | | | | | | | | | | | | |
| QY | 203 | ATQARPPPHASGPRRLGGERAWNHSVRAGVPLGLPAPGARRRGSGASRSLPKRPRR | 262 | | | | | | | | | | | | | | | | | |
| Db | 241 | GAAPERTPVGGGWAHFGTRTGPDSRDFCVVVSPPARPAEATSLGALSGTRHSHPSVG | 300 | | | | | | | | | | | | | | | | | |

QY 263 GAAPERTPVGOGSAHPCRTGPRGFCVSPARPAEATSLGALSCTRHSPSVG 322
Db 301 ROHAGPPSTRPPWDTPCPVPYAEATKHFYSSGDKQLRPSFLSSLRPSLTGARRL 360
QY 323 ROHAGPPSTRPPWDTPCPVPYAEATKHFYSSGDKQLRPSFLSSLRPSLTGARRL 382
Db 361 VETIFGSRPMPGTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPPLRAAVT 420
QY 383 VETIFGSRPMPGTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPPLRAAVT 442
Db 421 PAAGVCAREKPOGSVAAPAEEDTDPRLLVOLLRHSSPMQVYGFVRACLRRLVPPGLWGS 480
QY 443 PAAGVCAREKPOGSVAAPAEEDTDPRLLVOLLRHSSPMQVYGFVRACLRRLVPPGLWGS 502
Db 481 RHNERFLRNTKFIISLGKHAQLSLOELTWKMSVRDCAWLRSSRGVGCVPAAEHLREEI 540
QY 503 RHNERFLRNTKFIISLGKHAQLSLOELTWKMSVRDCAWLRSSRGVGCVPAAEHLREEI 562
Db 541 LAKFLHLMSSVYVVELLSRFSFYVTTTFOKNRLFYFKSVWSKLSQSIGIRHQLKRVOLRE 600
QY 563 LAKFLHLMSSVYVVELLSRFSFYVTTTFOKNRLFYFKSVWSKLSQSIGIRHQLKRVOLRE 622
Db 601 LSEAEVQHRREARPAALLTSRLRFTPKDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 660
QY 623 LSEAEVQHRREARPAALLTSRLRFTPKDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 682
Db 661 LFSVLNTERARRPGLLGASVLGLDDIHRAWRTFVLVRAQDPPPELYFVKVDVTGAYDTI 720
QY 683 LFSVLNTERARRPGLLGASVLGLDDIHRAWRTFVLVRAQDPPPELYFVKVDVTGAYDTI 742
Db 721 PODRLTEVIASIIKPONTYCVRRYAVVQKAAHGHVRKAFKSHVSTLDLOPYMRQFVAHL 780
QY 743 PODRLTEVIASIIKPONTYCVRRYAVVQKAAHGHVRKAFKSHVSTLDLOPYMRQFVAHL 802
Db 781 QETSPLDADVIEQSSSLNEASSGLFDVFLRFMCHHVRIRGKSVYOCQIGPOGSIILSTL 840
QY 803 QETSPLDADVIEQSSSLNEASSGLFDVFLRFMCHHVRIRGKSVYOCQIGPOGSIILSTL 862
Db 841 LCSICYGDMENKLFAGIRRDGLLLRLVDDELVTPLTHAKTFLRVLVIRGVPYGCYVNL 900
QY 863 LCSICYGDMENKLFAGIRRDGLLLRLVDDELVTPLTHAKTFLRVLVIRGVPYGCYVNL 922
Db 901 RKTVMNPFVEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVSQSDYSYARTSTRASLTF 960
QY 923 RKTVMNPFVEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVSQSDYSYARTSTRASLTF 982
Db 961 NRGEKAGNRMRKLFVLRKCHSLFDLDQVNSLOTCTNIYKILLQAVRFHACVLQLP 1020
QY 983 NRGEKAGNRMRKLFVLRKCHSLFDLDQVNSLOTCTNIYKILLQAVRFHACVLQLP 1042
Db 1021 FHOQWKNPTFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVOWLCHQAFLL 1080
QY 1043 FHOQWKNPTFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVOWLCHQAFLL 1102
Db 1081 KLTHRRVTVPLGSLRTAQTLRSKLPGLTTLTALEAAANPALSPDKFTILD 1132
QY 1103 KLTHRRVTVPLGSLRTAQTLRSKLPGLTTLTALEAAANPALSPDKFTILD 1154

RESULT 2
ID O14783 PRELIMINARY; PRT: 1132 AA.
AC O14783;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE TELOMERASE CATALYTIC SUBUNIT.
GN HEST2.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97433088.

RA MEYERSON M., COUNTER C.M., EATON E.N., ELLISEN L.W., STEINER P.,
RA CADDLE S.D., ZIAGRA L., BEIJERSBERG R.L., DAVIDOFF M.J., LIU Q.,
RA BACCHETTI S., HABER D.A., WEINBERG R.A.;
RL CELL 90:785-795(1997).
DR EMBL; AF018167; G2347129; -;
SQ SEQUENCE 1132 AA; 126937 MW; C1E5E2AF CRC32;
Query Match 98.1%; Score 8459; DB 4; Length 1132;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 1 MPRAPCRRAVRSLLRSHYREVLPPLATFVRRLPQCHRLVQRGDPAAFRALVAOCLVCPW 60
QY 23 MPRAPCRRAVRSLLRSHYREVLPPLATFVRRLPQCHRLVQRGDPAAFRALVAOCLVCPW 82
Db 61 DARPPPAAPSFQVSCCLKELVARVLQRLCERGAKNVLAFGFALLDARGGPPPEAFTTSVR 120
QY 83 DARPPPAAPSFQVSCCLKELVARVLQRLCERGAKNVLAFGFALLDARGGPPPEAFTTSVR 142
Db 121 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLARCALFVLVAPSCAYQVCGPPLYLQGA 180
QY 143 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLARCALFVLVAPSCAYQVCGPPLYLQGA 202
Db 181 ATQARPPPHASGPRRLGGERAWNHVSREAGVPLGLPAPGARBRGGSASRSLPLKPRR 240
QY 203 ATQARPPPHASGPRRLGGERAWNHVSREAGVPLGLPAPGARBRGGSASRSLPLKPRR 262
Db 241 GAAPERTPVGOGSAHPCRTGPRGFCVSPARPAEATSLGALSCTRHSPSVG 300
QY 263 GAAPERTPVGOGSAHPCRTGPRGFCVSPARPAEATSLGALSCTRHSPSVG 322
Db 301 ROHAGPPSTRPPWDTPCPVPYAEATKHFYSSGDKQLRPSFLSSLRPSLTGARRL 360
QY 323 ROHAGPPSTRPPWDTPCPVPYAEATKHFYSSGDKQLRPSFLSSLRPSLTGARRL 382
Db 361 VETIFGSRPMPGTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPPLRAAVT 420
QY 383 VETIFGSRPMPGTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPPLRAAVT 442
Db 421 PAAGVCAREKPOGSVAAPAEEDTDPRLLVOLLRHSSPMQVYGFVRACLRRLVPPGLWGS 480
QY 443 PAAGVCAREKPOGSVAAPAEEDTDPRLLVOLLRHSSPMQVYGFVRACLRRLVPPGLWGS 502
Db 481 RHNERFLRNTKFIISLGKHAQLSLOELTWKMSVRDCAWLRSSRGVGCVPAAEHLREEI 540
QY 503 RHNERFLRNTKFIISLGKHAQLSLOELTWKMSVRDCAWLRSSRGVGCVPAAEHLREEI 562
Db 541 LAKFLHLMSSVYVVELLSRFSFYVTTTFOKNRLFYFKSVWSKLSQSIGIRHQLKRVOLRE 600
QY 563 LAKFLHLMSSVYVVELLSRFSFYVTTTFOKNRLFYFKSVWSKLSQSIGIRHQLKRVOLRE 622
Db 601 LSEAEVQHRREARPAALLTSRLRFTPKDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 660
QY 623 LSEAEVQHRREARPAALLTSRLRFTPKDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 682
Db 661 LFSVLNTERARRPGLLGASVLGLDDIHRAWRTFVLVRAQDPPPELYFVKVDVTGAYDTI 720
QY 683 LFSVLNTERARRPGLLGASVLGLDDIHRAWRTFVLVRAQDPPPELYFVKVDVTGAYDTI 742
Db 721 PODRLTEVIASIIKPONTYCVRRYAVVQKAAHGHVRKAFKSHVSTLDLOPYMRQFVAHL 780
QY 743 PODRLTEVIASIIKPONTYCVRRYAVVQKAAHGHVRKAFKSHVSTLDLOPYMRQFVAHL 802
Db 781 QETSPLDADVIEQSSSLNEASSGLFDVFLRFMCHHVRIRGKSVYOCQIGPOGSIILSTL 840
QY 803 QETSPLDADVIEQSSSLNEASSGLFDVFLRFMCHHVRIRGKSVYOCQIGPOGSIILSTL 862
Db 841 LCSICYGDMENKLFAGIRRDGLLLRLVDDELVTPLTHAKTFLRVLVIRGVPYGCYVNL 900
QY 863 LCSICYGDMENKLFAGIRRDGLLLRLVDDELVTPLTHAKTFLRVLVIRGVPYGCYVNL 922
Db 901 RKTVMNPFVEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVSQSDYSYARTSTRASLTF 960
QY 923 RKTVMNPFVEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVSQSDYSYARTSTRASLTF 982
Db 961 NRGEKAGNRMRKLFVLRKCHSLFDLDQVNSLOTCTNIYKILLQAVRFHACVLQLP 1020
QY 983 NRGEKAGNRMRKLFVLRKCHSLFDLDQVNSLOTCTNIYKILLQAVRFHACVLQLP 1042
Db 1021 FHOQWKNPTFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVOWLCHQAFLL 1080
QY 1043 FHOQWKNPTFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVOWLCHQAFLL 1102
Db 1081 KLTHRRVTVPLGSLRTAQTLRSKLPGLTTLTALEAAANPALSPDKFTILD 1132
QY 1103 KLTHRRVTVPLGSLRTAQTLRSKLPGLTTLTALEAAANPALSPDKFTILD 1154

QY 923 RKTVMNFEDEALGGTAFAVQMPAHGLFPWCGLLDTRTLEQSDYSYARTSIRASLTF 982
Db 961 NRGFKAGNNRRKLFGLVRLKCHSLFLDLQVNSLQTVCTNIYKILLQAYRFAVYQLP 1020
QY 983 NRGFKAGNNRRKLFGLVRLKCHSLFLDLQVNSLQTVCTNIYKILLQAYRFAVYQLP 1042
Db 1021 FHOQWKNPTFFLRVSDTASLCSYLKAKNAGNSLGAAGPLPSEAWLWCHQAFLL 1080
QY 1043 FHOQWKNPTFFLRVSDTASLCSYLKAKNAGNSLGAAGPLPSEAWLWCHQAFLL 1102
Db 1081 KLTRHRYTVYVPLGLSLTAQTLQSKRLPGTTLTALEAANPALPDKFTILD 1132
QY 1103 KLTRHRYTVYVPLGLSLTAQTLQSKRLPGTTLTALEAANPALPDKFTILD 1154

RESULT 3
ID 070372 PRELIMINARY; PRT; 1122 AA.
AC 070372;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DE 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TELOMERASE REVERSE TRANSCRIPTASE.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA GREENBERG R.A., ALLSOPP R.C., CHIN L., MORIN G.B., DEPINHO R.A.;
RL ONCOGENE 0-0(1998).
DR EMBL: AF051911; G3005592; -
KW RNA-DIRECTED DNA POLYMERASE.
SQ SEQUENCE 1122 AA; 127977 MW; 222075d6 CRC32;

Query Match 59.8%; Score 5161; DB 11; Length 1122;
Best Local Similarity 62.7%; Pred. No. 0.00e+00;
Matches 718; Conservative 198; Mismatches 192; Indels 38; Gaps 24;

Db 1 MTRAPRCRAVRSLLRSYREVWPLATFVRLGPEGRRLVQPDGPKIYRTLVQAQCLVCMHW 60
QY 23 MPAPRCRAVRSLLRSYREVWPLATFVRLGPEGRRLVQPDGPKIYRTLVQAQCLVCMHW 82
Db 61 GSOPPPADLSFHOVSSLKELVARVVRQRCERNERNVLAFCFELLNEARGPPMAFTSVR 120
QY 83 DARPPAPRAFPGRVQSKELVARVLRQRCERGAKNVLAFCFELLNEARGPPMAFTSVR 142
Db 121 SYLPTVETLRVSGAMWLLSRVGDLLVYLLAHCALYLLVPPSCAYQVCGSPLYOICA 180
QY 143 SYLPTVTDALRGSGANGLLRVVDVLLHLLARCALFVLVAPSCAYQVCGSPLYOIGA 202
Db 181 TTDIWPVSASYPTRPVGRNFTWRLFLQIKSSROEAPKPLALPSRGTKRHLSLST 240
QY 203 ATQARPPPHAS-GPRRLGCE-RA-W--NH-SVR-EAGVPLGLPAPGARRRGGSSARS 253
Db 241 VPSAKKARCYPVPRVEGPHQVLPTRPSGKSWPSPARS-PEV-PT-AEKDLSS-KKVS 296
QY 254 LPLKPRRGAAPERTPTVQGSWAHPGRTGRPS-DRGFCVSPARPAGEATLEALS 312
Db 297 DLSSL-GSVCKKXPSSTLSLPPRONAFQLRP-FITRHLFYSRGDQERLNPSELLSN 354
QY 313 GTRHSHFVSGRQHAGPSTSRPPRPDTPCPVYATKHLFYSRGD-KEQRLSFELSS 371
Db 355 LQPNLTGARRLVEIFLGSRRPTSGPLCRTHLSRRYWMQRLPQQLVNHAEQCYVRL 414
QY 372 LRPSLTGARRLVETIFLGSRRPMTGPRRLPRLPQRYWMQRLPFLLELGNHQAQCPYVLL 431
Db 415 RSHCRFR---T--AN---Q--Q--VT--DALNTSPHLLMDLLKHSPPQVYGFRLACL 459
QY 432 KTHCPLEAAVTPAAGVCAREKPGQSAPEEEDTPRLVQLLRHQSPPQVYGFVRACL 491
Db 460 CKVVSASLWGRHNRHREFKLLKFFISLQYKGLSLQELMWMKVEDCHWLSPPGKDRV 519
QY 492 RRLVPPGLWGRHNRHREFKLLKFFISLQYKGLSLQELMWMKVEDCHWLSPPGKDRV 551

Db 520 PAAEHLRERILATFLFWLMDTVVOLLRSFFITESTFOKNRFLFFYRKSVWSKLSIGV 579
QY 552 PAAEHLRERILATFLFWLMDTVVOLLRSFFITESTFOKNRFLFFYRKSVWSKLSIGI 611
Db 580 ROHLERVLRELSEOEVRHODTWLAMPICRLRFIPKPNGLRPVNMYSWGTALGRKK 639
QY 612 ROHLKRVQLRELSEAEVRHREARPAALLSRFIPKPDGLRPIVNDYVVGARTFRREK 671
Db 640 QAQHTFORLKTFLSMLNYERTKPHLMSGSVLGMNDIYRTWRAFLVRLALDQTPRMFY 699
QY 672 RAERLTSRVKALFVNLNERARPPGLGASVLGDDHRAWRFTFLVRVAQDPPPELYFV 731
Db 700 KADVTGAYDAIPOGKLVVANNIRHSESTYCIQRYAVVRDSQGVHKSFRQVTTLS 759
QY 732 KVDVTGAYDTPQDLRTEVIAIIR-PONTYCVRYAVVQKAAHGHVYKAFKSHVSTLTD 790
Db 760 LQPMGQFLKHLQSDASALRNWVTEQSIEMNESSSLFDFLHLFRLHSHVVKIGRCYT 819
QY 791 LQPMRQFVAHLOET--SPLRDVAVIEQSSSLNEASGLDFELRMCCHAVIRKSYV 848
Db 820 QCQIGIPQSSSLTLLCSLCFCGMENKLFABEVQRDGLLLRFLVDDFLVTPHLDQAKTFLST 879
QY 849 QCQIGIPQSSSLTLLCSLCFCGMENKLFABEVQRDGLLLRFLVDDFLVTPHLDQAKTFLST 908
Db 880 LVHGVPEYGCMLNLTQTVNFPVPGTLGGAAPYQIPAHCLFPWCGLLLDTQILEVFCDY 939
QY 909 LVHGVPEYGCMLNLTQTVNFPVPGTLGGAAPYQIPAHCLFPWCGLLLDTQILEVFCDY 968
Db 940 SGVAOTSIXTSLTFSQSVKAGTKMRNKLKSLVRLKCHGLFLDLQVNSLQTVCTNIYKILL 999
QY 969 SSYARTSIRASLTFNRGFKAGNNRRKLFGLVRLKCHSLFLDLQVNSLQTVCTNIYKILL 1028
Db 1000 LQAYRFAVYQLPFPDQVRKNTLFFLGIISQASOACYALVKVKNPGMTL--K-AGSFP 1056
QY 1029 LQAYRFAVYQLPFPDQVRKNTLFFLGIISQASOACYALVKVKNPGMTL--K-AGSFP 1088
Db 1057 PERAHWCYQAFLLKLAHSAVYKLLGLPLRTAQKLLCRKLPEATMTILKAAADPALSTD 1116
QY 1089 SEAVWLCHQAFLLKLAHSAVYKLLGLPLRTAQKLLCRKLPEATMTILKAAADPALSTD 1148
Db 1117 FOTILD 1122
QY 1149 FKTILD 1154

RESULT 4
ID 013339 PRELIMINARY; PRT; 988 AA.
AC 013339;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TELOMERASE REVERSE TRANSCRIPTASE 1.
GN TET1
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972H-;
RX MEDLINE; 97400623.
RA NAKAMURA T.N., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H.,
RA LINGNER J., HARLEY C.B., CECI T.R.;
RL SCIENCE 277:955-959(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-972H-;
RA LYNE M., RAJANDREAM M.A., BARRELL B.G., VOLCKAERT G.;
RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF015783; G2340168; -
DR EMBL; AL022299; E1285360; -
DR PFAM; PF00078; ITC.
KW RNA-DIRECTED DNA POLYMERASE.
SQ SEQUENCE 988 AA; 116328 MW; 39C385A7 CRC32;

Db 651 FRKEMDYFRQKFOKTALEGGQYPTLFVLENDQNDLNAKTLIVAKORNYFKDNLL 710
QY 773 AHGHVRA-KFKSHVSTLT-DLQPTMRF-V-AHLQETSPRLDAVVIEQSSLSNEASSGLF 828
Db 711 QPVIN-ICQYNYINFNGKFKYQKGIPOGLCVSSILSFYATLESSLGFLRDESNPE 769
QY 829 DVLFREFCHHA-VRIKSKSVQCGIPQGSITLSCISLGYDMENK-L-F---AGTRD 882
Db 770 NPVNLMLRLTDYLLITTOENNAVLFIKLIINVSRENGFKFNKKLOTSPPLSPSKFAK 829
QY 883 G-LLRLVDDFLLVTPHLLTHAKTFLRVLRGVPEYGCVVNLKTVVFPVEDEALG- 937
Db 830 YGMDSEQINVDQYDQWIGISIDMKTLALPNIN-IRIEGILCTLNLMQTKASHWLK 888
QY 938 -GTAFVQMPAGH-L-F-PWCGLLDTRTLEQSDYSYVARTSIRASLTN-RGFKAGRNMR 993
Db 889 KKLKSL-MNNITHYFRKTTITTEFANKTKLFISSGYKYMOCAGE--YKDHFKKNLAM 945
QY 994 RKLGVRLKCHSLFDLQVNSLQTVCTNIYKILLQAYRHFACVLOLPHFQOQWKNPTF 1053
Db 946 SSMIDLEVKIIVSVTRA 963
QY 1054 FLRVISDTASLCYSILKA 1071

RESULT 7
ID Q35432 PRELIMINARY; PRT; 67 AA.
AC Q35432;

DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TELOMERASE CATALYTIC SUBUNIT (FRAGMENT).
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA DRISSE R., CLEVELAND J.L.;
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF029235; G2605903; --
FT NON-TER 1
FT NON-TER 67
SQ SEQUENCE 67 AA; 8368 MW; E2A06F2B CRC32;

Query Match 4.9%; Score 424; DB 11; Length 67;
Best Local Similarity 76.1%; Pred. No. 2.25e-57;
Matches 51; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

Db 1 FFYVTESTFQKRLFFYRKSVWSKLSQIGVROHLERVLRLSELSEVEVRHODTWLAMPIC 60
QY 582 FFYVTESTFQKRLFFYRKSVWSKLSQIGVROHLERVLRLSELSEVEVRHODTWLAMPIC 61

Db 61 RLREIPK 67
QY 642 RLREIPK 648

RESULT 8
ID Q06163 PRELIMINARY; PRT; 884 AA.
AC Q06163;

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE CHROMOSOME XII COSMID 8543.
GN L8543.12
OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMICOTINA; HEMIASCOMICETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RX MEDLINE; 97313267.
RA JOHNSTON M., HILLIER L., RILES L., ALBERMANN K., ANDRE B., ANSGORGE W.,

RA BENES V., BRUCKNER M., DELIUS H., DUBOIS E., DUSTERHOFT A.,
RA ENTIAN K.D., FLOETH M., GOFFEAU A., HEBLING U., HEUMANN K.,
RA HEUSS-NITTEL D., HILBERT H., HILGER F., KLEINE K., KOTTER P.,
RA LOUIS E.J., MESSENGUY F., MEWES H.W., MIOGA T., MOSTL D.,
RA MULLER-AUER S., NENTWICH U., OBERMAIER B., PIRAVANDI E., POHL T.M.,
RA PORTETELLE D., PURNELLE B., RECHMANN S., RIEGER M., RINKE M., ROSE M.,
RA SCHARPE M., SCHERENS B., SCHOLLER P., SCHWAGER C., SCHWARZ S.,
RA UNDERWOOD A.P., URRESTIAZU L.A., VANDENBOL M., VERHASSELT P.,
RA VIENDELS F., VOET M., VOLCKAERT G., VOSS H., WAMBUTT R., WEDLER E.,
RA WEDLER H., ZIMMERMANN F.K., ZOLLNER A., HANI J., HOHEISEL J.D.;
RN [2]
RP NATURE 387:0-0(0).
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RU DU 2.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA WATERSTON R.;
RL SUBMITTED (FEB-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA CHERRY J.M.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U20618; G662136; --
SQ SEQUENCE 884 AA; 102662 MW; 1A94320F CRC32;

Query Match 3.68%; Score 314; DB 3; Length 884;
Best Local Similarity 24.3%; Pred. No. 1.71e-35;
Matches 118; Conservative 122; Mismatches 202; Indels 44; Gaps 37;

Db 271 LSHLSRQSPKER-VLKFIIIVILQKLLPOEMFGSKKNGKIKLNLNLLSLPLNGYLPFDS 329
QY 470 LVQLLRQHSFPWQYGFVRACLRLLVPPGLWGRSHRERFLRTKFKFISLCKHAKLSQE 539
Db 330 LKLLRLKDFRWFIS-DIWFTHKFNENLNQALIC-FISWLFRLQIPKIIQTFYCTEIS 387
QY 530 LTWKMSVRDCAWLRSPGVCVPAAEHRLREILAKFLHLMSSVYVVELLSRFFYVTETT 589
Db 388 STVTIYF-RHDTWNLITPFIYVEPKTY-LVE-NNV-CRNHNSYLSNFENHSHKMLIPK 443
QY 590 FQXNRLFFYRKSVWSKLSQIGVROHLERVLRLSELSEVEVRHODTWLAMPIC 648
Db 444 KSNNEFRIIAIPCRGADDEEFTIYKENHK-NA-IQPTQKILEYLRNKRPTSF-TKIVSPT 500
QY 649 -PGLRPIVNDY-VVGARTFRREKRAERLTSVKALFSLVNTYERARRPGLLGASVLGLD 706
Db 501 QIADRI-KEFKQRLKFNVLPELYPMKFDVSKYCDSDIPRMECMR-ILKDALKNENGFF 558
QY 707 DI-HRAWRTVLRV-RA-QDPPPELYFVKVDVTGAYDTIPQ-DRLTEVIAIIPQNTYC 762
Db 559 VRSQY-FFNTINT-G-VLKLENV-VNASRVPKPY-ELYIDNVR-TVHLSNQDVIN---VV- 608
QY 763 VR-RYAVVOKAAGHVRAKFAKSHVSTLTDLPYMRQFVAHLQETSPRLDAVVIEQSSLSN 821
Db 609 E----M-EIF-K-T---ALWVEDKCYIREDLFGQSSLSAPIVDLVYDDLEFYSEFKAS 658
QY 822 EASSGLDFVLRFWCHHAVIRKGSVQCGIPQGSITLSCISLGYDM-E-NKLF-AG 878
Db 659 PSODTLILKLADDFLIISTDQOQVINIKKLAGGFGFKYAKAN-RDKILAVSSQSDDTV 717
QY 879 IRRDGLLLRLVDDFLVTPHLLTHAKTFLRVLRGVPEYGCVVNLKTVVFPVEDEALGG 938
Db 718 IQFCAM 723
QY 939 TAFVQM 944

RESULT 9
ID Q69118 PRELIMINARY; PRT; 296 AA.
AC Q69118;

01-NOV-1996 (TREMBLREL. 01, CREATED)
01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN (FRAGMENT).
OS HUMAN HERPESVIRUS TYPE 4.
OC VIRIDAE: DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE;
OC GAMMAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87284169.
RA FITZNER A.J., TSAI E.C., STROMINGER J.L., SPECK S.H.;
RL J. VIROL. 61:2902-2909(1987).
DR EMBL: M17294; G807646; -.
KW HYPOTHETICAL PROTEIN.
FT NON_TER 1
SQ SEQUENCE 296 AA; 31393 MW; D17CEF6F CRC32;

Query Match 1.7%; Score 147; DB 14; Length 296;
Best Local Similarity 38.6%; Pred. No. 7.18e-06;
Matches 39; Conservative 20; Mismatches 33; Indels 9; Gaps 8;

Db 91 GPPVTRSGAAQ-RTHRRPCCPSARNPCGPRTWRR-R-SGAQRGHPPPGAGQRPSPG 146
QY 194 GPPYQLGAATQARPPPHASG-PRR-R-LGCRARNHSVREAGVPLGLPAPGA-RRRGS 249
Db 147 TGGPAAAPGPTPAAPCGGGAAPVSGATPHPERGSGPAD 187
QY 250 ASRSLPLPKPRRGAAPPEPR-TPVGOGSWAHPGTRGPSD 289

RESULT 10
ID 036027 PRELIMINARY; PRT; 574 AA.
AC 036027;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE WISKOTT-ALDRICH SYNDROME PROTEIN HOMOLOG 1.
GN WSP1 OR SPAG4F10.15C
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JS21;
RA ZANKEL T.C., OW D.W.;
RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA CONNOR R., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WOOD V.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SIMILARITY: TO YEAST LAS17.
DR EMBL: AF038575; G2708709; -.
DR EMBL: Z98980; E339884; -.
FT DOMAIN 311 317 POLY-PRO.
FT DOMAIN 337 343 POLY-PRO.
FT DOMAIN 361 366 POLY-PRO.
FT VARIANT 248 248 L->V (IN STRAIN JS21).
SQ SEQUENCE 574 AA; 59605 MW; 9E84D37C CRC32;

Query Match 1.7%; Score 143; DB 3; Length 574;
Best Local Similarity 28.8%; Pred. No. 2.85e-05;
Matches 42; Conservative 41; Mismatches 53; Indels 10; Gaps 7;

Db 305 AANKRRPPPPPP-PSRR---NRG-KPPIGNGSSNSLPPPPPPPPSRNWAAG-SIPLPPQGR 358
QY 202 AATQARPPPHASGPRRLGCRARNHNSVREAGVPLGLPAPGARRRRGGSASRSLPLPKRPR 261
Db 359 SAPPPPPPSAPSGTOPPPGLSSSRAYSNNPP--APPAIPGRGAPALPPLGNASRTSTPP 416
QY 262 RGAAPERPTFVGOGSWAHP-GRTRGSGDRGFCVWSPARPAEATSTLEGALSGTRHSPS 320
Db 417 VPTPP-SLPPSPASPLPSPAPPSPLM 441

```

Query Match      1.68; Score 140; DB 2; Length 373;
Best Local Similarity 32.3%; Pred.No. 7.93e-05;
Matches 40; Conservative 27; Mismatches 46; Indels 11; Gaps 11;

Db 243 PRRAA-NRRRG-AAPAGGTAG-AAHG-HRRG-AARRARDRRRGGRRRGGQPRPG 297
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 215 PRRRLGGERAHNVSRGAVPLGLGPAGARRRGASASLPLKRRPRRGAAPERTPVG 274
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 L-RRARPPR-RGPPRRRPRAVGTSRPPRQAGG-AGH-R-AHRRTA-GRGGAARTHAOR 351
      : : | : : | : : | : : | : : | : : | : : | : : | : : | :
QY 275 QGSWAHPGTRGSDRGFCVVSPARPAEATSLEGALSGTRHSHPSVGRQHHAGPPSTR 334
      : : | : : | : : | : : | : : | : : | : : | : : | : : | :
Db 352 PGRP 355
      | | |
QY 335 PPRP 338
      | | |

RESULT 15
ID OA1935 PRELIMINARY; PRT; 585 AA.
AC OA1935;
DT 01-JAN-1998 (TREMBREL. 05, CREATED)
DT 01-JAN-1998 (TREMBREL. 05, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (TREMBREL. 05, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 60.2 KD PROTEIN.
GN GAMHAV.M6.
OS MURINE HERPESVIRUS 68.
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE; GAMMAHERPESVIRINAE
[1]
RN RN SEQUENCE FROM N.A.
RC STRAIN=WUMS;
RX MEDLINE; 97366649.
RA VIRGIN H.W. IV, LATREILLE P., WAMSLEY P., HALLSWORTH K., WECK K.E.,
RA DAL CANTO A.J., SPECK S.H.;
RA J. VIROL. 71:5894-5904(1997).
[2]
RN RN SEQUENCE FROM N.A.
RC STRAIN=WUMS;
RA LATREILLE P., WAMSLEY P., WATERSTON R.H.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U97553; G2317934; -
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 585 AA; 60160 MW; B35C72E5 CRC32;

Query Match      1.68; Score 135; DB 14; Length 585;
Best Local Similarity 27.6%; Pred. No. 4.23e-04;
Matches 45; Conservative 37; Mismatches 68; Indels 13; Gaps 11;

Db 342 PPP-ELGFGSPTAPASRAGARIPDLPLPLPSWGPDPRRPPPELPGSGSPTSPAPSRAG 400
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 195 PPLYQLGAANTQARPPPHASGPRRR-L-GCCRAHNSVREA-GVP-LGLPAPGARRGGSA 250
      : : | : : | : : | : : | : : | : : | : : | : : | : : | :
Db 401 ARIPDLPLPLPSWGPDPRRPPRPPPELGP-GSPT-SPASRAGARIPRSPRPPPELGP--G 456
      : : | : : | : : | : : | : : | : : | : : | : : | : : | :
QY 251 SRSLLPLPKR-PRRGAAPERTPVGGQSWAHPGTRGSDRGFCVVSPARPAEATSLEG 309
      : : | : : | : : | : : | : : | : : | : : | : : | : : | :
Db 457 SPREPRSRPP--PELPGSGFDLPLPLPSWG-PDPPTFPAPSR 496
      : : | : : | : : | : : | : : | : : | : : | : : | : : | :
QY 310 ALSGTRHSHPSVGRQHHAGPPSTRPPRPWDTPCPPVIAETKH 352
      : : | : : | : : | : : | : : | : : | : : | : : | : : | :

Search completed: Thu Dec 24 07:59:53 1998
Job time : 479 secs.

```

Search completed: Thu Dec 24 07:59:53 1998
Job time : 479 secs.

THIS PAGE BLANK (USPTO)

(TM)

| Result No. | Score | | Query Match | Length | DB | ID | Description | Pred. No. |
|------------|-------|-------|-------------|--------|--------|-----------------------|-------------|-----------|
| | Score | Match | | | | | | |
| 1 | 134 | 1.6 | 509 | 1 | P31060 | Predicted amino acid | 3 67e+01 | |
| 2 | 117 | 1.4 | 386 | 22 | W81864 | Predicted human NF-H | 5 36e+00 | |
| 3 | 118 | 1.4 | 399 | 1 | R23845 | Prepro form of human | 4 60e+00 | |
| 4 | 118 | 1.4 | 399 | 29 | W40193 | Mouse Op-2 protein. | 4 60e+00 | |
| 5 | 118 | 1.4 | 399 | 29 | W44314 | Mouse osteogenic prot | 4 60e+00 | |
| 6 | 118 | 1.4 | 399 | 16 | R35759 | Mouse osteogenic prot | 4 60e+00 | |
| 7 | 118 | 1.4 | 399 | 10 | R34938 | Mouse osteogenic prot | 4 60e+00 | |
| 8 | 118 | 1.4 | 399 | 9 | R33908 | Osteogenic protein mO | 4 60e+00 | |
| 9 | 118 | 1.4 | 399 | 9 | R46744 | Mouse osteogenic prot | 4 60e+00 | |
| 10 | 118 | 1.4 | 399 | 10 | R50201 | Mouse osteogenic prot | 4 60e+00 | |
| 11 | 118 | 1.4 | 399 | 9 | R47253 | Murine Op-2. | 4 60e+00 | |
| 12 | 118 | 1.4 | 399 | 19 | W00239 | Mouse osteogenic prot | 4 60e+00 | |
| 13 | 118 | 1.4 | 399 | 26 | W36856 | Full length sequence | 4 60e+00 | |
| 14 | 118 | 1.4 | 399 | 10 | R37974 | Murine Op-2. | 4 60e+00 | |
| 15 | 118 | 1.4 | 399 | 7 | R33411 | Murine pro-Op-2-PP. | 4 60e+00 | |
| 16 | 118 | 1.4 | 399 | 12 | R31646 | MOF2-PP prepro form m | 4 60e+00 | |
| 17 | 118 | 1.4 | 399 | 25 | W23144 | Bone morphogenic prot | 4 60e+00 | |
| 18 | 118 | 1.4 | 399 | 9 | R47293 | MOF2. | 4 60e+00 | |

b 37 pspstetessagtgasaprrptgdatpe-dstpvllrpw

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Db 3/ pspptetessagttagasaprrptgpdape-dstpvllrpgwgrsrpccpsttrpl

Db 37 pspstetessagttgasaprrtgdpatpe-dstpvlrrpwgrsrppcsttrpltns 95
|::: :|||: : ||| |::| :||| :||| :||| :||| :||| :||| :|||

```
QY 239 PAPGARRGGASRS-LPLPKRPR-RGAAPERTPVQGSWAHPGRTGRGSPD-RGFCVV 295
Db 96 tpppappedgrrpgagagnasrgprsggrrprprskappkwrk 140
QY 296 SPARPAEATSLGALSOTRHSHPVSGRQHAGPPSTSRP-PRDW 339

RESULT 2
ID W18664 standard; Protein; 386 AA.
AC W18664;
DE 24-JUL-1997 (first entry)
KW Framedgment human NF-H gene +1 frameshift mutant product.
KW Frameshift mutation product; GAGA motif; somatic mutation; diagnosis;
KW detection; antibody; probe; cancer; neoplasia; neurodegenerative;
KW Parkinson's; Alzheimer's disease; Pick's; Huntington's disease;
KW Down's syndrome; frontal lobe dementia; progressive supranuclear palsy;
KW PSP; amyotrophic lateral sclerosis; multiple sclerosis; MS;
KW cardiovascular; rheumatoid arthritis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT misc_difference 1..387
FT /note= "X corresponds to a stop codon in the
FT accompanying DNA file, T69796"
FT peptide
FT 129..135
FT /note= "antigenic peptide used fro antibody
FT production"
FT W09712992-A2.
FN 10-APR-1997.
PD 02-OCT-1996; IB1106.
PF 02-OCT-1995; GB-020080.
PR 11-JAN-1996; US-009832.
PR (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA (UYUO-) UNIV ROTTERDAM ERASMUS.
PA (UYUT-) UNIV STATE UTRCHTSM.
PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
DR WPI: 97-226235/20.
DR N-PSDB; T69796.
PT Use of mutant genes having frame:shift mutation(s) - for developing
PT prods. for the diagnosis, prevention and treatment of associated
PT diseases, e.g. cancer or neuro:degenerative disease
PT Claim 22; Fig 9; 123pp; English.
CC W18663 and W18664 are +2 and +1 frameshift mutations, respectively,
CC of a sequence comprising fragments of the coding sequence of the
CC human neurofilament subunit NF-H gene corresponding to nucleotides
CC 1-1162 of the wild-type NF-H gene. This region contains GAGAG motifs.
CC Frameshift mutants of the tau, ubiquitin, apolipoprotein E,
CC microtubule-associated protein 2 (MAP-2), neurofilament subunit L, M
CC and H and amyloid A4 genes are claimed. All these genes share a common
CC GAGAN motif (N= A, G, C or T), which is the site of common GA
CC dinucleotide deletion(s) that cause neurodegenerative disorders.
CC Antigenic peptides used for the production of antibodies, and small
CC nucleic acid sequences derived from frameshift mutants are used in the
CC diagnosis, prevention and treatment of cancer and neurodegenerative
CC diseases, e.g. Parkinson's disease, Alzheimer's disease, Down's
CC syndrome, frontal lobe dementia (Pick's disease), progressive
CC supranuclear palsy (PSP), amyotrophic lateral sclerosis, Huntington's
CC disease, multiple sclerosis, and other degenerative diseases such as
CC cardiovascular disease and rheumatoid arthritis.
SQ Sequence 386 AA;

Query Match 1.4%; Score 117; DB 22; Length 386;
Best Local Similarity 28.0%; Pred. No. 5.36e+00;
Matches 30; Conservative 26; Mismatches 50; Indels 1; Gaps 1;

Db 228 agglry-graviragpdraggaggrgatsatgagaparghrrapaprrirgpaarg 286
QY 232 AGVPLGLPAGARRGGASRS-LPLPKRPRGGAAPERTPVQGSWAHPGRTGRGSPDRG 291
Db 287 grggpragalragrggrgaepaeagagagvrlpaapppgrgga 333
QY 292 FCVVSPPARPAEATSLGALSOTRHSHPVSGRQHAGPPSTSRPRP 338
```

```
RESULT 3
ID R23846 standard; Protein; 399 AA.
AC R23846;
DE 05-NOV-1992 (first entry)
DE Prepro form of human osteogenic polypeptide-2.
KW Human osteogenic polypeptide-2; prepro sequence; implant;
KW endochondral bone formation; skeletal abnormality; dental abnormality;
KW non-union fracture; cartilage repair; osteoarthritis;
KW periodontal applcn.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cleavage_site 13..18
FT /label= cleavage site for signal peptide removal
FT cleavage_site 257..261
FT /label= cleavage site to form mature protein
FT protein 261..399
FT /label= mature protein 'hOP2-A'
FT protein 264..399
FT /label= truncated protein 'hOP2-P'
FT protein 267..399
FT /label= truncated protein 'hOP2-R'
FT /note= "I"
FT protein 240..399
FT /label= truncated protein 'hOP2-S'
FT region 303..399
FT /label= conserved six cysteine skeleton
FT region 297..399
FT /label= conserved seven cysteine skeleton
FN W09207073-A.
PN 30-APR-1992.
PD 18-OCT-1991; U07635.
PR 18-OCT-1990; US-599543.
PR (CREA-) CREATIVE BIOMOLECULES INC.
PA Kuberassampath T, Oppermann H, Ozkaynak E, Rueger DC;
DR WPI: 92-167153/20.
DR N-PSDB; N24522.
DR New osteogenic polypeptide(s) and their dimeric proteins - useful
PT in association with a matrix for bone reconstruction after
PT orthopaedic and plastic surgery
PT Disclosure; 51-53; 69pp; English.
CC The sequence encodes the prepro form of human osteogenic polypeptide-2
CC (hOP2-PP). The protein is useful as a subunit of osteogenic
CC proteins capable of endochondral bone formation, and allogenic and
CC xenogenic implants in mammals. When implanted with a matrix, the
CC polypeptide locally induces the full developmental cascade of
CC endochondral bone formation and bone marrow differentiation. This
CC can be used for optimal predictable bone formation to correct, eg
CC acquired or congenital craniofacial and other skeletal or dental
CC anomalies, and to treat non-union fractures. The polypeptide also
CC has dental and periodontal applicns., and may be used for cartilage
CC repair, eg in osteoarthritis treatment. See also W09105802.
SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 1; Length 399;
Best Local Similarity 39.7%; Pred. No. 4.60e+00;
Matches 25; Conservative 13; Mismatches 19; Indels 6; Gaps 6;

Db 12 glalcalggggpgrpp-qc-pqrrlg-ardrdvq-reilavlgpgr-prprappaas 66
QY 194 GPPYQLCAATQA-RPPPHASGPRRLRCERANHNSVREAGVPLGLPAGARRGGASR 252
Db 67 rlp 69
QY 253 SLP 255

RESULT 4
ID W40193 standard; Protein; 399 AA.
AC W40193;
DE 08-JUN-1998 (first entry)
DE Mouse OP-2 protein.
KW Osteogenic protein; OP-2; bone morphogenic protein; OP/BMP family;
KW chronic renal failure; renal therapeutic agent; disease; diabetes;
```


PR 23-FEB-1989; US-315342.
 PR 17-OCT-1989; US-422699.
 PR 17-OCT-1989; US-422613.
 PR 22-FEB-1990; US-483913.
 PR 20-AUG-1990; US-569920.
 PR 07-SEP-1990; US-579865.
 PR 18-OCT-1990; US-599543.
 PR 21-NOV-1990; US-616374.
 PR 04-DEC-1990; US-621849.
 PR 04-DEC-1990; US-621988.
 PR 22-FEB-1991; US-660162.
 PR 20-DEC-1991; US-810560.
 PR 28-JAN-1992; US-827052.
 PR 21-FEB-1992; US-841646.
 PR 01-NOV-1993; US-147023.
 PR (STYK) STRYKER CORP
 PI Kuberasampath T, Oppermann H, Ozkaynak E, Rueger DC;
 DR WPI: 96-010159/01.
 DR N-PSDB; 702599.
 PT Antibodies with osteogenic protein binding specificity - used in
 PT purification of osteogenic proteins, and as antigenic proteins
 PS Disclosure; Column 139-42; 129pp; English.
 CC This sequence represents the murine osteogenic protein, mOP-2.
 CC has homology with proteins in the TGF-beta superfamily. Fragments
 CC of the mOP-2 protein can be used in the production of dimeric
 CC peptides which may be used in the generation of antibodies with
 CC binding specificities for osteogenic proteins. The antibodies are
 CC capable of binding specifically to an epitope of the osteogenic protein
 CC and may be used in purification protocols. Osteogenic proteins, such as
 CC mOP-2, may be used in an implantable osteogenic device which allows
 CC predictable bone formation to correct acquired and congenital
 CC craniofacial and other skeletal or dental anomalies. They may be used to
 CC induce local endochondral bone formation in non-union fractures and in
 CC other clinical applications including dental and periodontal applications
 CC where bone formation is required. Other potential applications include
 CC cartilage repair, e.g. in the treatment of osteoarthritis.
 SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 16; Length 399;
 Best Local Similarity 38.7%; Pred. No. 4.60e+00;
 Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 glalcalggghgppp-htc-pqrrlg-arerrdmqreillavlgpr-prpraqpaar 67
 QY 194 GPLYQLGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRRGSASRS 253
 Db 68 qp 69
 QY 254 LP 255

RESULT 7
 ID R54938 standard; Protein; 399 AA.
 AC R54938;
 DT 15-OCT-1994 (first entry)
 DE Osteogenic protein mOP2-PP.
 KW Morphogenic protein; mOP-2-PP; OP-2; mOP2; mOP-2;
 KW Tissue morphogenesis; osteogenic protein.
 OS Mus sp.
 PN W09410203-A.
 PD 11-MAY-1994.
 PF 02-NOV-1993; U10520.
 PR 03-NOV-1992; US-971091.
 PR 04-MAR-1993; US-029335.
 PR 31-MAR-1993; US-040510.
 PA (CREA-) CREATIVE BIOMOLECULES INC.
 PI Cohen CM, Kuberasampath T, Oppermann H, Ozkaynak E;
 PI Pang RHL, Rueger DC;
 DR WPI: 94-167392/20.
 DR P-PSDB; 065394.
 PT A morphogenically active protein MOP-3 - for inducing tissue
 PT morphogenesis in mammals
 PS Disclosure; Page 134-136; 164pp; English.

CC A novel mouse morphogenic protein, OP3, has the sequence given in
 CC R54934, and is encoded by cDNA of sequence 065390. cDNA and protein
 CC sequences were also provided for human osteogenic protein OP1
 CC (065391, R54935), mouse OP1 (065392, R54936), human OP2 (065393,
 CC R54937) and mouse OP2 (065394, R54938), as well as the genomic DNA
 CC sequence of human OP2 (065395). Generic sequences given in R54939-
 CC 40 accommodate homologies between OP1, OP2, OP3 and other morphogen
 CC protein family members.
 SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 10; Length 399;
 Best Local Similarity 38.7%; Pred. No. 4.60e+00;
 Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 glalcalggghgppp-htc-pqrrlg-arerrdmqreillavlgpr-prpraqpaar 67
 QY 194 GPLYQLGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRRGSASRS 253
 Db 68 qp 69
 QY 254 LP 255

RESULT 8
 ID R33908 standard; Protein; 399 AA.
 AC R33908;
 DT 21-JUL-1993 (first entry)
 DE Mouse osteogenic protein 2 (mOP-2).
 KW Bone; loss; increase; fracture; post-menopausal; senile;
 KW osteoporosis; hyperparathyroidism; skeletal microstructure defects;
 KW chronic renal failure; kidney disease; osteomalacia; vitamin D;
 KW deficiency-induced osteopenia, osteoporosis; Paget's disease;
 KW bone mass; imbalance; resorption; formation; dialysis; calcium;
 KW phosphate; metabolism; murine.
 OS Mus musculus.
 FH Key
 FT Location/Qualifiers
 FT protein 261..399
 FT region 18..260
 FT /note= "mature protein"
 FT /note= "pro region, cleaved to yield mature,
 FT active protein"
 FT region 298..299
 FT /note= "conserved seven cysteine skeleton"
 PN W09305751-A.
 PD 01-APR-1993.
 PF 28-AUG-1992; U07432.
 PR 30-AUG-1991; US-752764.
 PR 30-AUG-1991; US-752857.
 PR 30-AUG-1991; US-752861.
 PR 31-JUL-1992; US-923780.
 PA (CREA-) CREATIVE BIOMOLECULES INC.
 PI Cohen CM, Kuberasampath T, Oppermann H, Ozkaynak E,
 PI Pang RHL, Rueger DC, Smart JE;
 DR WPI: 93-117208/14.
 PT Use of morphogenic or in-vivo morphogenic-stimulating agent - to
 PT prevent bone loss or increase, used for treating bone fractures,
 PT post-menopausal or senile osteoporosis, hyperparathyroidism etc.
 PS Disclosure; Page 122-123; 162pp; English.
 CC The sequence is that of mouse osteogenic protein 2 (mOP-2) a
 CC a morphogenically active protein which may be used as part of a
 CC method for treating a bone fracture or a disease which causes or
 CC results in bone fractures or other defects in skeletal
 CC microstructure. Such diseases include chronic renal failure and
 CC other kidney diseases, osteomalacia, vitamin D deficiency-induced
 CC osteopenia or osteoporosis, postmenopausal or senile osteoporosis,
 CC hyperparathyroidism and Paget's disease. The methods can be used for
 CC protecting individuals at risk for loss of bone mass such as
 CC postmenopausal females, aged individuals and individuals undergoing
 CC dialysis. The loss of bone mass may result from an imbalance in bone
 CC resorption or bone formation, an imbalance of calcium or phosphate
 CC metabolism, a vitamin D imbalance or be nutritionally or hormonally
 CC induced.
 SQ Sequence 399 AA;

| | | | | | | |
|--------|---|-----------------------|--------|---------------------|--------|----------------|
| OS | Mus musculus. | Query Match | 1.4%; | Score 118; | DB 10; | Length 399; |
| PN | WO9406420-A. | Best Local Similarity | 38.7%; | Pred. No. 4.60e+00; | | |
| PP | 31-MAR-1993. | Matches | 24; | Conservative | 12; | Mismatches 22; |
| PP | 15-SEP-1993. | | | | | Indels 4; |
| PP | 15-SEP-1992; | | | | | Gaps |
| PP | 04-MAR-1993; | | | | | |
| PP | 31-MAR-1993; | | | | | |
| PP | (CREA-) CREATIVE BIOMOLECULES INC. | | | | | |
| PI | Charette MF, Cohen CM, Kuberasampath T, Oppermann H; | | | | | |
| PI | Ozkaynak E, Pang RHL, Rueger DC, Smart JE; | | | | | |
| PI | WPI: 94-118121/14. | | | | | |
| DR | N-PSDB: Q45119. | | | | | |
| DR | Maintaining integrity of gastrointestinal lining using a | | | | | |
| PFT | morphogen (stimulant) - for treating or preventing ulceration, | | | | | |
| PFT | also to inhibit endothelial cell proliferation and reduce side | | | | | |
| PFT | effects of cancer therapy. | | | | | |
| PPS | Claim 35-36: Page 115-117: 151pp; English. | | | | | |
| CCC | Morphogens comprising an amino acid sequence sharing at least | | | | | |
| CCC | 70% homology with OP-1, OP-2, BMP2, BMP3(fx), Vg1(fx), Vgr(fx), | | | | | |
| CCC | DPP(fx), GDF-1(fx), 60A(fx) and at least 80% homology with | | | | | |
| CCC | BMP5(fx) and BMP6(fx) are useful for maintaining the integrity of | | | | | |
| CCC | the gastrointestinal tract luminal lining in a mammal, including | | | | | |
| CCC | (1) limiting epithelial cell proliferation, (2) inhibiting ulcerative | | | | | |
| CCC | lesion formation, (3) inhibiting inflammation normally associated | | | | | |
| CCC | with ulcerative diseases, and/or (4) stimulating the repair of | | | | | |
| CCC | ulcerative lesions and the regeneration of the luminal tissue. | | | | | |
| SSQ | Sequence 399 AA; | | | | | |
| Db | 12 glalcalggghprpp-htc-pqrrlg-arerdmqreilavlglpgr-prpragpaaar 67 | | | | | |
| QY | 194 GPPLQLGNAQTARPPPHASGPRRLGCCERANWNSVRVGLPLGAPCARRGGSASRS 253 | | | | | |
| Db | 68 qp 69 | | | | | |
| QY | 254 LP 255 | | | | | |
| RESULT | 11 | | | | | |
| ID | R47253 standard; Protein; 399 AA. | | | | | |
| AD | R47253; | | | | | |
| DT | 15-AUG-1994 (first entry) | | | | | |
| DE | MOP2. | | | | | |
| KW | Human; hippocampus; osteogenic protein; OP-1; subunit; dimer; | | | | | |
| KW | morphogenic activity; cysteine; morphogen; family; pro-region; | | | | | |
| KW | complex; soluble; aqueous solvent; therapeutic composition; | | | | | |
| KW | symptom-alleviating; co-factor; antibody; diagnosis; assay; | | | | | |
| KW | quantitate; mature. | | | | | |
| OS | Mus musculus. | | | | | |
| PN | WO9403600-A. | | | | | |
| PP | 17-FEB-1994. | | | | | |
| PP | 29-JUL-1993; | | | | | |
| PP | 31-JUL-1992; | | | | | |
| PP | 04-MAR-1993; | | | | | |
| PP | 31-MAR-1993; | | | | | |
| PI | (CREA-) CREATIVE BIOMOLECULES INC. | | | | | |
| PI | Jones WK, Kuberasampath T, Oppermann H, Ozkaynak E; | | | | | |
| PI | Rueger DC, Tucker RF; | | | | | |
| PI | WPI: 94-065689/08. | | | | | |
| DR | N-PSDB: Q36201. | | | | | |
| DR | Morphogenic protein soluble complex - for regeneration of tissue | | | | | |
| PFT | in mammals and diagnosing tissue disorders | | | | | |
| PFT | Claim 3; Page 75-77: 120pp; English. | | | | | |
| CCC | This sequence represents the murine derived protein, osteogenic | | | | | |
| CCC | protein, mOP-2. The mature OP-2 protein was used as at | | | | | |
| CCC | least one subunit in the dimeric protein of the invention. This | | | | | |
| CCC | dimeric protein comprises a pair of protein subunits which are | | | | | |
| CCC | associated to give a structure with morphogenic activity. Each | | | | | |
| CCC | subunit comprises more than 100 amino acids having a pattern of | | | | | |

CC cysteine residues characteristic of the morphogen family. Each
CC subunit comprises a mature form of a subunit of a member of the
CC morphogen family, non-covalently complexed with a peptide comprising
CC a pro-region of a morphogenic family member, to form a complex more
CC soluble in aqueous solvents than the uncomplexed subunits. The
CC dimeric protein is useful in a therapeutic composition, pref. also
CC containing a symptom-alleviating cofactor. The protein and also
CC corresponding antibody may be used in diagnostic assays, eg. to
CC quantitate the amount of mature and soluble forms of morphogenic
CC proteins produced.
SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 9; Length 399;

Best Local Similarity 38.7%; Pred. No. 4.60e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 glalcalggghgprpp-htc-pqrrlg-arerrdmqrellavlgpgr-prpragpaaar 67
QY 194 GPPYLQGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRRGSASRS 253
Db 68 qp 69
QY 254 LP 255

RESULT 12
ID W00239 standard; Protein; 399 AA.

AC W00239;
DT 21-NOV-1996 (first entry)
DE Mouse osteogenic protein OP-2
KW Morphogen; osteogenic protein; dentine; tooth decay; caries;
KW morphogenesis; odontoblast; OP-2.
OS Mus musculus.

FH Key Location/Qualifiers
FT peptide 1..17
FT /label= Sig_peptide
FT region 18..260
FT /label= Pro_region
FT protein 261..399
FT /label= Mat_protein
FT domain 288..399
FT /label= 7-Cys_C-terminal_domain

PN W09626737-AL.
PD 06-SEP-1996.
PF 14-FEB-1996; U02169.
PR 01-MAR-1995; US-396930.
PA (CREA-) CREATIVE BIOMOLECULES INC.
PI Charette MF, Rutherford RB;
DR WPI: 96-412583/41.
DR N-PSDB; T33444.
PT Use of morphogen(s), e.g. osteogenic proteins, on dental surfaces
PT - for inducing dentine morphogenesis, desensitising teeth or sealing
PT tooth cavities
PS Disclosure: Page 61-63: 106pp; English.
CC Mouse embryo full-length osteogenic protein OP-2 (W00239)
CC includes a pro-sequence and the morphogenically active mature
CC protein sequence (see also W00224) that includes a 7-Cys C-terminal
CC domain. OP-2 can be expressed from intact or truncated cDNA
CC (T33444) in prokaryotic or eukaryotic host cells. Mature OP-2 and
CC other morphogens (see also W00221-35), partic. human OP-1 (W00221),
CC can be used to induce dentine morphogenesis, to seal dental
CC cavities and to desensitise teeth to pressure and/or temp.
SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 19; Length 399;

Best Local Similarity 38.7%; Pred. No. 4.60e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 glalcalggghgprpp-htc-pqrrlg-arerrdmqrellavlgpgr-prpragpaaar 67
QY 194 GPPYLQGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRRGSASRS 253
Db 68 qp 69

QY 254 LP 255

RESULT 13

ID W36856 standard; Protein; 399 AA.

AC W36856;
DT 10-MAR-1998 (first entry)
DE Full length sequence of mouse osteogenic protein 2 (hop-2).
KW Mouse osteogenic protein; OP; OP-2; morphogen; morphogenic protein;
KW embryogenesis; organ maintenance; tissue-specific morphogenesis;
KW arthritis; emphysema; osteoporosis; cirrhosis.
OS Mus sp.

FH Key Location/Qualifiers
FT Region 18..260
FT /note= "pro region which is cleaved to yield the
FT mature morphogenically active protein"

FT Protein 261..399
FT /note= "mature protein"
FT Region 297..399
FT /note= "conserved 7 Cys skeleton"

US5650276-A.

PD 22-JUL-1997.
PF 20-JUL-1994; 278729.
PR 28-AUG-1992; US-938021.
PR 30-AUG-1991; US-752764.
PR 30-AUG-1991; US-752861.
PA (CREA-) CREATIVE BIOMOLECULES INC.
PI Cohen CM, Kuberasampath T, Oppermann H, Ozkaynak E;
PI Pang RH, Rueger DC, Smart JE;
DR WPI: 97-384865/35.
DR N-PSDB; T97882.

PT Screening for compounds which modulate morphogen expression - by
PT incubating in the presence of epithelial cells which contain a
PT cellular gene for morphogenic protein expression
PT Disclosure: Columns 67-70: 49pp; English.
PS The present sequence represents a mouse osteogenic protein-2 (mOP-2).
CC OP-2 proteins are a group of morphogenically active proteins. Morphogens
CC are inactive when reduced, but are active as oxidised homodimers and when
CC oxidised with other morphogens (e.g. W36853-62). Comparison of the amino
CC acid sequences of these morphogens has identified a consensus 6-7
CC cysteine motif at the C-terminal. Morphogenic proteins such as OP-2 play
CC an important role, not only in embryogenesis, but also in tissue and
CC organ maintenance and repair in mammals. They induce a developmental
CC cascade of tissue-specific morphogenesis in a mammal. A novel method is
CC described for screening a candidate compound for the ability to modulate
CC expression of a cellular gene encoding a naturally occurring morphogenic
CC protein. The candidate compound is incubated with epithelial cells which
CC express the cellular gene, and after a period of time the epithelial
CC cells are assayed for the presence of or the amount of the protein
CC expressed by the cellular gene. A change in the level of the
CC morphogenic protein relative to the level in the epithelial cells in the
CC absence of the candidate compound is indicative of the ability of the
CC compound to modulate expression of the cellular gene. The method can be
CC used to identify compounds which can increase or decrease morphogen
CC production or levels. Such compounds can be used in the treatment of,
CC e.g. arthritis, emphysema, osteoporosis, kidney disease, lung diseases,
CC cardiomyopathy, and cirrhosis of the liver.

SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 26; Length 399;

Best Local Similarity 38.7%; Pred. No. 4.60e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 glalcalggghgprpp-htc-pqrrlg-arerrdmqrellavlgpgr-prpragpaaar 67
QY 194 GPPYLQGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRRGSASRS 253
Db 68 qp 69

QY 254 LP 255

RESULT 14

ID R57974 standard; Protein; 399 AA.
 AC R57974; 1994 (first entry)
 DT 11-OCT-1994 (first entry)
 DE Murine OP-2.
 KW OP-1; OP-2; CBMP2; Vg1(fx); Vgr(fx); DPP(fx);
 KW GDF-1(fx); 60A(fx); BMP3(fx); BMP5(fx); BMP6(fx);
 KW tooth socket; alveolus; osteogenic protein; morphogen;
 KW morphogenic protein; periodontal tissue; regeneration;
 KW tooth implant; integration; inhibition.
 OS Mus musculus.
 PN WO9406399-A.
 PD 31-MAR-1994.
 PF 15-SEP-1993; U08742.
 PR 15-SEP-1992; US-945285.
 PR 04-MAR-1993; US-029335.
 PR 31-MAR-1993; US-040510.
 PA (CREA-) KUBERASAMPATH T, OPPERMANN H, OZKAYNAK E;
 PI Pang RHL, Rueger DC, Smart JE;
 PI WPI; 94-118107/14.
 DR N-PSDB; 067314.
 DT Morphogen-induced periodontal tissue regeneration - used in
 PT integrating as implanted tooth in tooth socket or to inhibit
 PT tissue loss associated with periodontal disease or injury
 PS Claim 28-29; Page 100-103; 132pp; English.
 CC Morphogens comprising an amino acid sequence sharing at least
 CC 70% homology with Op-1, Op-2, CBMP2, Vg1(fx), Vgr(fx), DPP(fx),
 CC GDF-1(fx), 60A(fx) and at least 80% homology with BMP3(fx),
 CC BMP5(fx) and BMP6(fx) are useful for integrating an implanted
 CC tooth in a tooth socket and for inhibiting tissue loss associated
 CC with periodontal disease or injury.
 SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 10; Length 399;
 Best Local Similarity 38.7%; Pred. NO. 4.60e+00;
 Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Query Match 1.4%; Score 118; DB 10; Length 399;
 Best Local Similarity 38.7%; Pred. NO. 4.60e+00;
 Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 glalcalggghgrpp-htc-pqrrlg-arerrdmqreilavlgpgr-prpraqpaaar 67
 QY 194 GPLYQLGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRRGSASRS 253
 Db 68 qp 69
 QY 254 LP 255

RESULT 15

ID R33411 standard; Protein; 399 AA.
 AC R33411;
 DT 15-JUL-1993 (first entry)
 DE Murine pro-Op-2-pp.
 KW morphogenic; osteogenic protein; developmental cascade; mOP-2;
 KW mouse; inflammation; anti-inflammatory; Transforming Growth Factor;
 KW TGF-beta super-family; hippocampus.
 OS Mus.
 FH Key Location/Qualifiers
 FT cds 93..1292
 FT /*tag= a
 FT /*note= "mOP-2-pp"
 FT mat_peptide 873..1289
 FT /*tag= b
 FT /*note= "contains conserved 7 cysteine skeleton"
 PN WO9304692-A.
 PD 18-MAR-1993.
 PF 28-AUG-1992; U07358.
 PR 30-AUG-1991; US-752764.
 PR 30-AUG-1991; US-752861.
 PR 30-AUG-1991; US-753059.
 PA (CREA-) CREATIVE BIOMOLECULES INC.
 PI Cohen CM, Kuberampath T, Oppermann H, Ozkaynak E;
 PI Pang RHL, Rueger DC, Smart JE;
 PI WPI; 93-100652/12.
 DR

DR N-PSDB; Q38736.
 DT Morphogen-induced modulation of inflammatory response - and
 PT resulting tissue damage, e.g. in autoimmune diseases, diabetes,
 PT asthma, ischemia reperfusion injury, etc.
 PS Claim 26; Page 126-127; 165pp; English.
 CC Osteogenic protein (OP)-2 is a preferred morphogen
 CC for use in treating tissue damage in e.g. inflammatory disease, and
 CC autoimmune disease, arthritis, psoriasis, dermatitis, diabetes, and
 CC emphysema. Proteins having at least 70% homology with OP-2 amino
 CC acid sequences can also be used. See R33401 for mature mOP-2.
 SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 7; Length 399;
 Best Local Similarity 38.7%; Pred. NO. 4.60e+00;
 Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 glalcalggghgrpp-htc-pqrrlg-arerrdmqreilavlgpgr-prpraqpaaar 67
 QY 194 GPLYQLGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRRGSASRS 253
 Db 68 qp 69
 QY 254 LP 255

Search completed: Thu Dec 24 07:42:27 1998
 Job time : 339 secs.

THIS PAGE BLANK (USPTO)

WIPERH
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Dec 28 07:52:22 1998; MasPar time 58.82 Seconds

Tabular output not generated. 742.293 Million cell updates/sec

Title: >US-08-951-733-20

Description: (1-1154) from US08951733.pep

Perfect Score: 8624

Sequence: 1 HASGQRVLLRTWEALAPAT.....TALEAANPALPSDFKTLTD 1154

Scoring table: PAM 150

Gap 11

Searched: 320804 seqs, 37834689 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: a-pending

1: P9 2: U60 3: U7 4: U80 5: U81 6: U82 7: U83 8: U84 9: U85

10: U86 11: U87 12: U88 13: U89 14: U90 15: U91 16: NEWP

17: NEWU6 18: NEWU8 19: NEWU9

Statistics: Mean 40.462; Variance 196.348; scale 0.206

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Query | ID | Description | Pred. No. |
|------------|-------|-------|---------|------------|------------------------|-----------|
| 1 | 8624 | 100.0 | 1154 13 | US-08-951- | Sequence 20, Applicati | 0.00e+00 |
| 2 | 8607 | 99.8 | 1189 18 | US-08-974- | Sequence 613, Applicat | 0.00e+00 |
| 3 | 8607 | 99.8 | 1189 13 | US-08-911- | Sequence 34, Applicati | 0.00e+00 |
| 4 | 8607 | 99.8 | 1189 13 | US-08-912- | Sequence 325, Applicat | 0.00e+00 |
| 5 | 8607 | 99.8 | 1200 18 | US-08-974- | Sequence 612, Applicat | 0.00e+00 |
| 6 | 8607 | 99.8 | 1200 13 | US-08-912- | Sequence 324, Applicat | 0.00e+00 |
| 7 | 8607 | 99.8 | 1200 13 | US-08-911- | Sequence 33, Applicati | 0.00e+00 |
| 8 | 8607 | 99.8 | 1285 18 | US-08-974- | Sequence 600, Applicat | 0.00e+00 |
| 9 | 8607 | 99.8 | 1285 13 | US-08-911- | Sequence 32, Applicati | 0.00e+00 |
| 10 | 8607 | 99.8 | 1285 13 | US-08-912- | Sequence 314, Applicat | 0.00e+00 |
| 11 | 8590 | 99.6 | 1407 13 | US-08-911- | Sequence 55, Applicati | 0.00e+00 |
| 12 | 8590 | 99.6 | 1407 18 | US-08-974- | Sequence 628, Applicat | 0.00e+00 |
| 13 | 8590 | 99.6 | 1407 13 | US-08-912- | Sequence 334, Applicat | 0.00e+00 |
| 14 | 8465 | 98.2 | 1132 18 | US-08-974- | Sequence 2, Applicatio | 0.00e+00 |
| 15 | 8465 | 98.2 | 1132 13 | US-08-912- | Sequence 2, Applicatio | 0.00e+00 |
| 16 | 8465 | 98.2 | 1132 13 | US-08-911- | Sequence 2, Applicatio | 0.00e+00 |
| 17 | 8465 | 98.2 | 1132 14 | US-09-042- | Sequence 3, Applicatio | 0.00e+00 |
| 18 | 8465 | 98.2 | 1132 14 | US-09-052- | Sequence 2, Applicatio | 0.00e+00 |
| 19 | 8465 | 98.2 | 1132 12 | US-08-954- | Sequence 225, Applicat | 0.00e+00 |
| 20 | 8465 | 98.2 | 1154 13 | US-08-912- | Sequence 323, Applicat | 0.00e+00 |
| 21 | 8465 | 98.2 | 1154 18 | US-08-974- | Sequence 611, Applicat | 0.00e+00 |

22 8465 98.2 1154 13 US-08-911- Sequence 45, Applicati 0.00e+00
 23 8459 98.1 1132 14 US-09-026- Sequence 3, Applicati 0.00e+00
 24 8459 98.0 1132 18 US-08-974- Sequence 344, Applicat 0.00e+00
 25 7113 82.5 949 13 US-08-951- Sequence 14, Applicati 0.00e+00
 26 7113 82.5 949 13 US-08-912- Sequence 5, Applicati 0.00e+00
 27 5742 66.6 807 13 US-08-911- Sequence 5, Applicati 0.00e+00
 28 5742 66.6 807 18 US-08-974- Sequence 5, Applicati 0.00e+00
 29 5334 61.9 1003 18 US-08-974- Sequence 336, Applicat 0.00e+00
 30 5334 61.9 1003 12 US-08-854- Sequence 217, Applicat 0.00e+00
 31 5334 61.9 1003 12 US-08-851- Sequence 217, Applicat 0.00e+00
 32 5161 59.8 1122 14 US-09-042- Sequence 2, Applicati 0.00e+00
 33 4070 47.2 545 13 US-08-911- Sequence 47, Applicati 0.00e+00
 34 2458 28.5 564 18 US-08-974- Sequence 267, Applicat 2.37e-199
 35 2458 28.5 564 12 US-08-854- Sequence 101, Applicat 2.37e-199
 36 2458 28.5 564 12 US-08-846- Sequence 101, Applicat 2.37e-199
 37 2458 28.5 564 12 US-08-851- Sequence 101, Applicat 2.37e-199
 38 2226 25.8 538 13 US-08-912- Sequence 316, Applicat 6.64e-179
 39 2226 25.8 538 18 US-08-974- Sequence 602, Applicat 6.64e-179
 40 2226 25.8 538 13 US-08-911- Sequence 46, Applicati 6.64e-179
 41 2129 24.7 514 13 US-08-911- Sequence 41, Applicati 2.30e-170
 42 2129 24.7 514 13 US-08-912- Sequence 319, Applicat 2.30e-170
 43 2129 24.7 514 18 US-08-974- Sequence 605, Applicat 2.30e-170
 44 2085 24.2 530 13 US-08-912- Sequence 317, Applicat 1.71e-166
 45 2085 24.2 530 18 US-08-974- Sequence 603, Applicat 1.71e-166

ALIGNMENTS

RESULT 1
 ID US-08-951-733-20 STANDARD; PRT; 1154 AA.
 XX
 AC xxxxxx
 XX
 DT
 XX
 XX
 DE
 XX

Sequence 20, Application US/08951733

Sequence 20, Application US/08951733

GENERAL INFORMATION:

APPLICANT: Harrington, Lea A.

APPLICANT: Robinson, Murray O.

TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Angen Inc.

STREET: One Angen Center Drive

CITY: Thousand Oaks

STATE: CA

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/951,733

FILING DATE: 16-OCT-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/873,039

FILING DATE: 11-JUN-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/751,189

FILING DATE: 15-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Oleski, Nancy A.

REGISTRATION NUMBER: 34,688

REFERENCE/DOCKET NUMBER: A-433B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (805) 447-6504

TELEFAX: (805) 499-8011

INFORMATION FOR SEQ ID NO: 20:

CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1154 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: unknown
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
SQ SEQUENCE 1154 AA; 129326 MW; 6842246 CN;

Query Match 100.0%; Score 8624; DB 13; Length 1154;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 HASGQCVLLRTWEALAPATPAMPAPRCRAVRSLLRSYREVLPPLATFVRRLGPGQWRL 60
QY 1 HASGQCVLLRTWEALAPATPAMPAPRCRAVRSLLRSYREVLPPLATFVRRLGPGQWRL 60
Db 61 VORGDPAAFRALVAOCLVCPWDARPPAPSPFROVSCLELKEIVARVQLRCLCERGAKNVLA 120
QY 61 VORGDPAAFRALVAOCLVCPWDARPPAPSPFROVSCLELKEIVARVQLRCLCERGAKNVLA 120
Db 121 FGFALLDARGGPPPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGGDDVLLHLLARCAL 180
QY 121 FGFALLDARGGPPPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGGDDVLLHLLARCAL 180
Db 181 FVLVAPSCAYQVCGPLYQLGATQARPPPHASGPRRLGCRANVHVSREAGVPLGLPA 240
QY 181 FVLVAPSCAYQVCGPLYQLGATQARPPPHASGPRRLGCRANVHVSREAGVPLGLPA 240
Db 241 PGARRGGASRSLLPKPRGAPPEPRTPVGGSWAHGPGTRGSPDRGFCVVSARP 300
QY 241 PGARRGGASRSLLPKPRGAPPEPRTPVGGSWAHGPGTRGSPDRGFCVVSARP 300
Db 301 ABEATLEGALSGTRHSHPSVGRQHAGPPSTSRPPRWDTPCPPIYAEKHFYSSGDK 360
QY 301 ABEATLEGALSGTRHSHPSVGRQHAGPPSTSRPPRWDTPCPPIYAEKHFYSSGDK 360
Db 361 EQLRPSFLLSLRPSLTGARLLVEIFLGSRPWPGTTPRLRLPQRYWQMRPLFLELIG 420
QY 361 EQLRPSFLLSLRPSLTGARLLVEIFLGSRPWPGTTPRLRLPQRYWQMRPLFLELIG 420
Db 421 NHAQCPYGVLLKTHCPPLRAAVTPAAGVCAREKPGQSVAAPEEDTPRLVOLLRQHSP 480
QY 421 NHAQCPYGVLLKTHCPPLRAAVTPAAGVCAREKPGQSVAAPEEDTPRLVOLLRQHSP 480
Db 481 WQYGFVRACLRLVPPGLWGSRHNERFLRNTKFFISLGKHAQLSLOELTWKMSVRDCA 540
QY 481 WQYGFVRACLRLVPPGLWGSRHNERFLRNTKFFISLGKHAQLSLOELTWKMSVRDCA 540
Db 541 WLRRSPGVGCVPAAEHLREEILAKFLHWMYSYVVELLSRFYTTTFQKNRLFYRK 600
QY 541 WLRRSPGVGCVPAAEHLREEILAKFLHWMYSYVVELLSRFYTTTFQKNRLFYRK 600
Db 601 SVWSKLQSIGIRQHLKRVOLRELSAEVROHREARPAALLTSRLRIPKPDGLRPIVNDY 660
QY 601 SVWSKLQSIGIRQHLKRVOLRELSAEVROHREARPAALLTSRLRIPKPDGLRPIVNDY 660
Db 661 VVGARTFREKRAERLTSRKALFVSNLYERARRPCLLGASVGLDDIHRAWRTFVLVR 720
QY 661 VVGARTFREKRAERLTSRKALFVSNLYERARRPCLLGASVGLDDIHRAWRTFVLVR 720
Db 721 AQDPPPELVFVKVDYTGAYDTIPQDRLETVIASIKPQNTYCVRRYAVVQKAAGHVRKA 780
QY 721 AQDPPPELVFVKVDYTGAYDTIPQDRLETVIASIKPQNTYCVRRYAVVQKAAGHVRKA 780
Db 781 FKSHVSTLTLQPYMRQFVAHQETSPLRDVAVVIQSSSLNEASSLDFVFLRMCHAV 840
QY 781 FKSHVSTLTLQPYMRQFVAHQETSPLRDVAVVIQSSSLNEASSLDFVFLRMCHAV 840
Db 841 RIRGKSYVQCQIPQGSILSTLCSLCYGDMEKNKLFAGIRRDGLLRLVDDFLLVTPHLT 900
QY 841 RIRGKSYVQCQIPQGSILSTLCSLCYGDMEKNKLFAGIRRDGLLRLVDDFLLVTPHLT 900
Db 901 HAKTFLRLTVRGPEYGCVVNLRKTVVNFVEDEALGTAFAVQMPAHGLFPWCGLLLDTR 960

SEQUENCE 613, Application US/08974549
Sequence 613, Application US/08974549
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morlin, Gregg B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 726
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312

CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph Ted
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-0026100S
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 613:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1189 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: Protein
CC LOCATION: 1..1189
CC OTHER INFORMATION: /note= "fusion protein composed of
CC OTHER INFORMATION: melittin signal sequence and full length
CC OTHER INFORMATION: htrt protein"
SQ SEQUENCE 1189 AA; 133179 MW; 7256545 CN;

Query Match 99.8%; Score 8607; DB 18; Length 1189;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 37 ASTORCVLLRTWEALPATPAMPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRV 96
|||
Qy 2 ASGQRCVLLRTWEALPATPAMPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRV 61

Db 97 QRGDPAAFRALVAOCLVCVWDARPPAPAFSROVSKLKLVARVLQRCERGAKNVLA 156
|||
Qy 62 QRGDPAAFRALVAOCLVCVWDARPPAPAFSROVSKLKLVARVLQRCERGAKNVLA 121

Db 157 GFALLDGARGPPFAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALF 216
|||
Qy 122 GFALLDGARGPPFAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALF 181

Db 217 VLVA PCSAYQVCGPPLYLQGAATQARPPPHASGPRRLGRCERAWNHSVREAGVPLGLPAP 276
|||
Qy 182 VLVA PCSAYQVCGPPLYLQGAATQARPPPHASGPRRLGRCERAWNHSVREAGVPLGLPAP 241

Db 277 GARRRGASRSPLPKRRPRGAPEPRTVPGGSAHAPGRTGRPSDRGFCVVSAPRA 336
|||
Qy 242 GARRRGASRSPLPKRRPRGAPEPRTVPGGSAHAPGRTGRPSDRGFCVVSAPRA 301

Db 337 EEATSLGALSSTHSHPSVGRQHGAGPPSTSRPRPDWTPCPVYATKHFVLSGDK 396
|||
Qy 302 EEATSLGALSSTHSHPSVGRQHGAGPPSTSRPRPDWTPCPVYATKHFVLSGDK 361

Db 397 QLRSFLLSLRLPGLTARRLVETIFLGSRPWMPGTFRRLPRLPQRYWQMRPLFLELLGN 456
|||
Qy 362 QLRSFLLSLRLPGLTARRLVETIFLGSRPWMPGTFRRLPRLPQRYWQMRPLFLELLGN 421

Db 457 HAQCPYGVLLKTHCPPLRAAVTPAAGVCAREKPGQSAAPPEEDTDPRRLVQLLRQHSPPW 516
|||
Qy 422 HAQCPYGVLLKTHCPPLRAAVTPAAGVCAREKPGQSAAPPEEDTDPRRLVQLLRQHSPPW 481

Db 517 QVYGFVRACLRLRVLPPGLWGRHNRERFLNTKKFISLGHAKLSLOELTWKMSVRDCAW 576
|||

Qy 482 QVYGFVRACLRLRVLPPGLWGRHNRERFLNTKKFISLGHAKLSLOELTWKMSVRDCAW 541
Db 577 LRRSPGVGCVPAAEHRLREILAKFLHMLSVYVWELLRSFFVYVETTFQKNRLFYRKS 636
|||
Qy 542 LRRSPGVGCVPAAEHRLREILAKFLHMLSVYVWELLRSFFVYVETTFQKNRLFYRKS 601
|||
Db 637 VWSKLSIGIRQHLKRVQLRELSEAEVROHREARPAALLTSRLRFIPKPDGLRIPVNDYV 696
|||
Qy 602 VWSKLSIGIRQHLKRVQLRELSEAEVROHREARPAALLTSRLRFIPKPDGLRIPVNDYV 661
|||
Db 697 VGARTFREKRAERLTSRVKALFSLVNYERARRPGLLGASVLGLDDTHRAWRFTVLVRVA 756
|||
Qy 662 VGARTFREKRAERLTSRVKALFSLVNYERARRPGLLGASVLGLDDTHRAWRFTVLVRVA 721
|||
Db 757 QDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIKPNQTYCVRRYAVVQKAAGHVRKAF 816
|||
Qy 722 QDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIKPNQTYCVRRYAVVQKAAGHVRKAF 781
|||
Db 817 KSHVSTLTDLQPMRQFVAHLQETSPLRDAVVEQSSSLNEASSGLDFVFLRFMCHHAVR 876
|||
Qy 782 KSHVSTLTDLQPMRQFVAHLQETSPLRDAVVEQSSSLNEASSGLDFVFLRFMCHHAVR 841
|||
Db 877 IRGKSYVQCGIPOGSIILSLCVCYGDMMENKLFAGIRRDGLLLRLVDDFLVTPHLTH 936
|||
Qy 842 IRGKSYVQCGIPOGSIILSLCVCYGDMMENKLFAGIRRDGLLLRLVDDFLVTPHLTH 901
|||
Db 937 AKTFLRTLVRGVPPEYGCVVNLKRTVYVNFVEDEALGCTAFVQMPAHGLFPWCGLLLDTRT 996
|||
Qy 902 AKTFLRTLVRGVPPEYGCVVNLKRTVYVNFVEDEALGCTAFVQMPAHGLFPWCGLLLDTRT 961
|||
Db 997 LEVQSDYSSTARTSIRASLTFRNGFAGRNMRKLFGLVKHLSLFLDLQVNSLQVCT 1056
|||
Qy 962 LEVQSDYSSTARTSIRASLTFRNGFAGRNMRKLFGLVKHLSLFLDLQVNSLQVCT 1021
|||
Db 1057 NIKKILLQAYRFHACVQLQPFHQOVWKNPTFFLRLVSDTASLCYSTLKAKNAGMSLGAK 1116
|||
Qy 1022 NIKKILLQAYRFHACVQLQPFHQOVWKNPTFFLRLVSDTASLCYSTLKAKNAGMSLGAK 1081
|||
Db 1117 GAAGPLPSEAVQWLCHQAFLLKLRHRTVYVPLGLSLRTAQTLQSLRKLPGTTLTALFAAA 1176
|||
Qy 1082 GAAGPLPSEAVQWLCHQAFLLKLRHRTVYVPLGLSLRTAQTLQSLRKLPGTTLTALFAAA 1141
|||
Db 1177 NPALPSDFKILD 1189
|||
Qy 1142 NPALPSDFKILD 1154
|||

RESULT 3
ID US-08-911-312-34 STANDARD; PRT; 1189 AA.
XX
AC xxxxxx
XX
DT
XX
XX
XX
DE Sequence 34, Application US/08911312
-XX
CC Sequence 34, Application US/08911312
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase
CC NUMBER OF SEQUENCES: 170
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA

CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA: US/08/911,312
CC APPLICATION NUMBER: US/08/911,312
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Einhorn, Gregory P.
CC REGISTRATION NUMBER: 38,440
CC REFERENCE/DOCKET NUMBER: 015389-002500US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 34:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1189 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1189 AA; 133179 MW; 7256545 CN;

Query Match 99.8%; Score 8607; DB 13; Length 1189;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

37 ASTQRCVLLRTWEALAPATPAMPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRV 96
11
2 ASGORCVLLRTWEALAPATPAMPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRV 61
97 QRGDPAFRALVAOCLVCVPWDAPPPAAPSPFQVSCLELVARVQLRCERGAKNVIAF 156
62 QRGDPAFRALVAOCLVCVPWDAPPPAAPSPFQVSCLELVARVQLRCERGAKNVIAF 121
157 GFALLDGARGGPPPEAFVTSVRSYLPNTVTDALRGSGAWGLLLRRVDDVLLHLLARCALF 216
122 GFALLDGARGGPPPEAFVTSVRSYLPNTVTDALRGSGAWGLLLRRVDDVLLHLLARCALF 181
217 VLVAAPCAVQVCGPPLYLQGAATQARPSPHAGSPRRRLGCRANVHNSVREAGVPLGLPAP 276
182 VLVAAPCAVQVCGPPLYLQGAATQARPSPHAGSPRRRLGCRANVHNSVREAGVPLGLPAP 241
277 GARRGGASRSLLPLPKRRGAAPPEPTEPVGGGSAHAPGRTGRPSDRGFCVYSPARPA 336
242 GARRGGASRSLLPLPKRRGAAPPEPTEPVGGGSAHAPGRTGRPSDRGFCVYSPARPA 301
337 EEATSLGALSCTRHSHPSVGRQHHAGPPSTSRPPRWDTPCPVVAETKHFVSSGDKE 396
11

QY 302 EEATSLGALSCTRHSHPSVGRQHHAGPPSTSRPPRWDTPCPVVAETKHFVSSGDKE 361
Db 397 QLRPSFLLSLRPSLATGARRLVETIFLGSPPMPGPRRLPRLPQRYWQMRPLFLELLGN 456
QY 362 QLRPSFLLSLRPSLATGARRLVETIFLGSPPMPGPRRLPRLPQRYWQMRPLFLELLGN 421
Db 457 HAOCPTYGVLLKTHCPLRAAVTPAAGVCAREKPOGSAVAAPEEEDTDPRRLVQLLRQHSSPW 516
QY 422 HAOCPTYGVLLKTHCPLRAAVTPAAGVCAREKPOGSAVAAPEEEDTDPRRLVQLLRQHSSPW 481
Db 517 QVTGVRACILRLVPPGLWGRHNERFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 576
QY 482 QVTGVRACILRLVPPGLWGRHNERFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 541
Db 577 LRRSPGVGCVPAAEHRLREELAKFLHMLSVYVVELLSRFFVYVTTTFOKNRLFYRKS 636
QY 542 LRRSPGVGCVPAAEHRLREELAKFLHMLSVYVVELLSRFFVYVTTTFOKNRLFYRKS 601
Db 637 WSKLASIGIRQHLKRVQLRELSEAEVQRHREARPAALLTSRLRFIPKPDGLRPIVNNMYV 696
QY 602 WSKLASIGIRQHLKRVQLRELSEAEVQRHREARPAALLTSRLRFIPKPDGLRPIVNNMYV 661
Db 697 VGARTFRERKRAERLTSRVKALFSLVNLNERARRPGLLGASVLGLDDHRAWRTFVLVRVA 756
QY 662 VGARTFRERKRAERLTSRVKALFSLVNLNERARRPGLLGASVLGLDDHRAWRTFVLVRVA 721
Db 757 QDPPELYFYKVDVTGAYDTIPQDRLTEVIASIKPONTYCVRRYAVVQKAAHGHVKAF 816
QY 722 QDPPELYFYKVDVTGAYDTIPQDRLTEVIASIKPONTYCVRRYAVVQKAAHGHVKAF 781
Db 817 KSHVSTLTDLPYMRQFVAHLQETSPLRDAVIEQSSSLNEASSGLFDVFLRFMCHHAVR 876
QY 782 KSHVSTLTDLPYMRQFVAHLQETSPLRDAVIEQSSSLNEASSGLFDVFLRFMCHHAVR 841
Db 877 IRKSYVQCOGIPQGSITLSTLSCYGDENKLFAGIRBDGILLRLVDDFLVTPHLTH 936
QY 842 IRKSYVQCOGIPQGSITLSTLSCYGDENKLFAGIRBDGILLRLVDDFLVTPHLTH 901
Db 937 AKTFLRTLVRGVPYEGCVVNLKRTVNVFVVEDEALGGTAFAVQMPAHGLFPWCGLLDTRT 996
QY 902 AKTFLRTLVRGVPYEGCVVNLKRTVNVFVVEDEALGGTAFAVQMPAHGLFPWCGLLDTRT 961
Db 997 LEVQSDYSYARTSIRASLTFRNGFKAGRNRRKLFGLVRLKCHSLFLDLQVNSLQTVCT 1056
QY 962 LEVQSDYSYARTSIRASLTFRNGFKAGRNRRKLFGLVRLKCHSLFLDLQVNSLQTVCT 1021
Db 1057 NIYKILLQAYRFHACVLOLPPHQVWKNPTFFFLRVISDTASLCYSTLKAKNAGMSLGAK 1116
QY 1022 NIYKILLQAYRFHACVLOLPPHQVWKNPTFFFLRVISDTASLCYSTLKAKNAGMSLGAK 1081
Db 1117 GAAGPLPSEAVQWLCHQAFLLKLTTRHRTVYVPLIGSLRTAQTLQSLRKLPGTTTLTAEAAA 1176
QY 1082 GAAGPLPSEAVQWLCHQAFLLKLTTRHRTVYVPLIGSLRTAQTLQSLRKLPGTTTLTAEAAA 1141
Db 1177 NPALPSDFKTILD 1189
QY 1142 NPALPSDFKTILD 1154

RESULT 4
ID US-08-912-951-325 STANDARD; PRT; 1189 AA.
XX
AC xxxxxx
XX
DT
XX
Sequence 325, Application US/08912951
XX
Sequence 325, Application US/08912951
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC TITLE OF INVENTION: THERAPEUTIC METHODS
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Apple, Randolph T.

CC REGISTRATION NUMBER: 36,429

CC REFERENCE/DOCKET NUMBER: 015389-002600US

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (415) 576-0200

CC TELEFAX: (415) 576-0300

CC INFORMATION FOR SEQ ID NO: 325:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 1189 amino acids

CC TYPE: amino acid

CC STRANDEDNESS:

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

CC SEQUENCE 1189 AA; 133179 MW; 7256545 CN;

Query Match 99.88; Score 8607; DB 13; Length 1189;

Best Local Similarity 99.98; Pred. No. 0.00e+00;

Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

37 ASTQRCVLLRTWEALPATPAMPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQGRNV 96

2 ASGQRCVLLRTWEALPATPAMPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQGRNV 61

97 QRGDPAAFRALVAQCLVCPWMDARPPAPPSFRQVSKLKVLRQLRCERGAKNYLA 156

62 QRGDPAAFRALVAQCLVCPWMDARPPAPPSFRQVSKLKVLRQLRCERGAKNYLA 121

157 GFALLDGARGGPEAFTTSVRSYLPNTVTDALRGSGANGLLLRVGGDVLVHLLARCALF 216

QY 122 GFALLDGARGGPEAFTTSVRSYLPNTVTDALRGSGANGLLLRVGGDVLVHLLARCALF 181
DB 217 VLVPASCAVQCGPPPLYQLGAATQARPPHASPRLRLGCERAWNSHREAGVPLGLPAP 276
QY 182 VLVPASCAVQCGPPPLYQLGAATQARPPHASPRLRLGCERAWNSHREAGVPLGLPAP 241
DB 277 GARRGGASRSRLPLKPRRGAAPERTPVQGGWAHPGRTRGSDRGFCVVSARPA 336
QY 242 GARRGGASRSRLPLKPRRGAAPERTPVQGGWAHPGRTRGSDRGFCVVSARPA 301
DB 337 EEATSLLEGALSGTRHSHPSVGRQHHAGPSTSRPPRPWDTCPVPVYAEKHFYSSGDKE 396
QY 302 EEATSLLEGALSGTRHSHPSVGRQHHAGPSTSRPPRPWDTCPVPVYAEKHFYSSGDKE 361
DB 397 QLRPSFLLSLRPSLTGARLVETIFLGRPMWPGTPRRLPLRQRYWQMRPLFLELLGN 456
QY 362 QLRPSFLLSLRPSLTGARLVETIFLGRPMWPGTPRRLPLRQRYWQMRPLFLELLGN 421
DB 457 HAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPOGSGVAAPPEEDTDPRRLVQLLRQHSPPW 516
QY 422 HAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPOGSGVAAPPEEDTDPRRLVQLLRQHSPPW 481
DB 517 QVYGFVRACLRRLVPPGLAGSRHNERREFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 576
QY 482 QVYGFVRACLRRLVPPGLAGSRHNERREFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 541
DB 577 LRSPGVCVPAAEHLREEILAKFLHLMSSVYVVELLSFFVYVTTTFOKNLFFYRKS 636
QY 542 LRSPGVCVPAAEHLREEILAKFLHLMSSVYVVELLSFFVYVTTTFOKNLFFYRKS 601
DB 637 VWSKLSIGIRQHLKRVQLRELSAEVROHREARPAALLTSRLRFIPKPDGLRPIVNDYV 696
QY 602 VWSKLSIGIRQHLKRVQLRELSAEVROHREARPAALLTSRLRFIPKPDGLRPIVNDYV 661
DB 697 VGARTFERREKRAERLTSRVKALFSVLNYSERARPPGLLGASVLGLDDIHRARWTFVLVRVA 756
QY 662 VGARTFERREKRAERLTSRVKALFSVLNYSERARPPGLLGASVLGLDDIHRARWTFVLVRVA 721
DB 757 QDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIKPNQTYCVRRYAVVQKAAHGHVRKAF 816
QY 722 QDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIKPNQTYCVRRYAVVQKAAHGHVRKAF 791
DB 817 KSHVSTLTLQPYMROFVAHLQETSPLRDAVTEQSSSLNEASSGLFDVFLRPMCHHAVR 876
QY 782 KSHVSTLTLQPYMROFVAHLQETSPLRDAVTEQSSSLNEASSGLFDVFLRPMCHHAVR 841
DB 877 IRGKSYVQCGIPIQSGSILSTLLCSLCYGDWENKLFAGIRDDGLLRLVDDFLVTPHLTH 936
QY 842 IRGKSYVQCGIPIQSGSILSTLLCSLCYGDWENKLFAGIRDDGLLRLVDDFLVTPHLTH 901
DB 937 AKTFLETLVRGVPPEYGCVVNLRKTVVNFVPEDEALGTAFAVQMPAHGLFPWCGLLDTRT 996
QY 902 AKTFLETLVRGVPPEYGCVVNLRKTVVNFVPEDEALGTAFAVQMPAHGLFPWCGLLDTRT 961
DB 997 LEVQSDYSYARTSIRASLTFNRGFKAGRMRRKLFGLVLRKLSHSLFLDLQVNSLQTVCT 1056
QY 962 LEVQSDYSYARTSIRASLTFNRGFKAGRMRRKLFGLVLRKLSHSLFLDLQVNSLQTVCT 1021
DB 1057 NIYKILLQAYRPHACVLOLPFHQQVWKNTPELRYISDTASLCYSILKAKNAGMSLGAK 1116
QY 1022 NIYKILLQAYRPHACVLOLPFHQQVWKNTPELRYISDTASLCYSILKAKNAGMSLGAK 1081
DB 1117 GAAGPLPSEAVQWMLCHQAFLLKLTNRHVTVVPLGLSLRTAQTOLSKRLPOTTTTALAEEA 1176
QY 1082 GAAGPLPSEAVQWMLCHQAFLLKLTNRHVTVVPLGLSLRTAQTOLSKRLPOTTTTALAEEA 1141
DB 1177 NPALPSDFKTILD 1189
QY 1142 NPALPSDFKTILD 1154
RESULT 5
ID US-08-974-549-612 STANDARD; PRT; 1200 AA.

XX
AC
XX
DT
XX
XX
DE
XX
XX
Sequence 612, Application US/08974549
XX
Sequence 612, Application US/08974549
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Human Telomerase Catalytic Subunit
CC NUMBER OF SEQUENCES: 726
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/974,549
CC FILING DATE: 19-NOV-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph Ted
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-0026100S
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 612:
CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 1200 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: Protein
CC LOCATION: 1..1200
CC OTHER INFORMATION: /note= "fusion protein composed of His6
CC OTHER INFORMATION: and Anti-Xpress tags, enterokinase
CC OTHER INFORMATION: cleavage site and full length HTRT
CC OTHER INFORMATION: protein"
CC SQ SEQUENCE 1200 AA; 134322 MW; 7387257 CW;

Query Match 99.8%; Score 8607; DB 18; Length 1200;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 48 ASTQRCVLLTWEALAPATPAMPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQGRWLV 107
QY 2 ASQRCVLLTWEALAPATPAMPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQGRWLV 61

Db 108 QRGDPAAFRALVAQCLVCPWDARPPAPSPFQVSCLELVARVLRQRCERGAKNVLA 167
QY 62 QRGDPAAFRALVAQCLVCPWDARPPAPSPFQVSCLELVARVLRQRCERGAKNVLA 121

Db 168 GFALLDGARGGPEAFTTSVRSYLPNTVDALRGSGAWGLLRVGGDDVLVHLLARCALF 227
QY 122 GFALLDGARGGPEAFTTSVRSYLPNTVDALRGSGAWGLLRVGGDDVLVHLLARCALF 181

Db 228 VLVAPSCAYQVCGPPLYQLGAATQARPPPHASGPRRLCERAWNSHVREAGVPLGLPAP 287
QY 182 VLVAPSCAYQVCGPPLYQLGAATQARPPPHASGPRRLCERAWNSHVREAGVPLGLPAP 241

Db 288 GARRRGGSASRLPLKPRRGAAPERTPVQGSWAHPGTRGSDRGFCVVSARPA 347
QY 242 GARRRGGSASRLPLKPRRGAAPERTPVQGSWAHPGTRGSDRGFCVVSARPA 301

Db 348 EEATSLLEGALSGTRHSHPSVGRQHAGPSTSRPPPMWTPCPVYAEKHFLYSSGDKE 407
QY 302 EEATSLLEGALSGTRHSHPSVGRQHAGPSTSRPPPMWTPCPVYAEKHFLYSSGDKE 361

Db 408 QLRSFLLSRLPSLTGARLVETIFLGSRPWMPGTPRRLPLRPQRYWQMRPLFLELLGN 467
QY 362 QLRSFLLSRLPSLTGARLVETIFLGSRPWMPGTPRRLPLRPQRYWQMRPLFLELLGN 421

Db 468 HAOCYPGVLLKTHCPLRAAVTPAAGVCAKEKPGQGSVAAPPEEDTDPRLLVOLLRQHS 527
QY 422 HAOCYPGVLLKTHCPLRAAVTPAAGVCAKEKPGQGSVAAPPEEDTDPRLLVOLLRQHS 481

Db 528 QYGFVRACLRLRVLPPGLWGSRHNERFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 587
QY 482 QYGFVRACLRLRVLPPGLWGSRHNERFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 541

Db 588 LRSPGVGCVPAEAHRLREILAKFLHMLMSVYVVELLSFFYVTTTQKNRLLFFYRKS 647
QY 542 LRSPGVGCVPAEAHRLREILAKFLHMLMSVYVVELLSFFYVTTTQKNRLLFFYRKS 601

Db 648 VMSKLSIGIRQHLKRVQLRELSEAEVROHREARPAALLTSRLRFIPKPGCLRPVNDYV 707
QY 602 VMSKLSIGIRQHLKRVQLRELSEAEVROHREARPAALLTSRLRFIPKPGCLRPVNDYV 661

Db 708 VGARTFRREKRAERLT SRVKALFVSNLYERARRPGLLGASVLGLDDIHRARTFVLVRVA 767
QY 662 VGARTFRREKRAERLT SRVKALFVSNLYERARRPGLLGASVLGLDDIHRARTFVLVRVA 721

Db 768 QDPPPELYFKVDVTGAYDTIPQDRLTEVIASIKPONTYCVRRYAVVQKAAHGHVRKAF 827
QY 722 QDPPPELYFKVDVTGAYDTIPQDRLTEVIASIKPONTYCVRRYAVVQKAAHGHVRKAF 781

Db 828 KSHVSTLTDLQPYMROFVAHLQETSPLRDAVTEQSSSLNEASSGLDFVLRPMCHAVR 887
QY 782 KSHVSTLTDLQPYMROFVAHLQETSPLRDAVTEQSSSLNEASSGLDFVLRPMCHAVR 841

Db 888 IRKSVVQOGIPQGSILSTLLCSLCYGD MENKLFAGIRRDG LLLRLVDDFLVTPHLTH 947
QY 842 IRKSVVQOGIPQGSILSTLLCSLCYGD MENKLFAGIRRDG LLLRLVDDFLVTPHLTH 901
Db 948 AKTFRLTVRGVPEYGVNLRKTVNFPVEDEALGTAFOVQPAHGLFPWCGLLDTRT 1007
QY 902 AKTFRLTVRGVPEYGVNLRKTVNFPVEDEALGTAFOVQPAHGLFPWCGLLDTRT 961
Db 1008 LEVQSDYSYARTSIRASLTFNRFGRKAGNRNRKLFVGLRKCHSLFLDLQVNSLQTVCT 1067
QY 962 LEVQSDYSYARTSIRASLTFNRFGRKAGNRNRKLFVGLRKCHSLFLDLQVNSLQTVCT 1021
Db 1068 NIKYKILLQAYRHACVQLPFPHQVWKNPTFFLRVISTDASLCYSILKAKNAGMSLGAK 1127
QY 1022 NIKYKILLQAYRHACVQLPFPHQVWKNPTFFLRVISTDASLCYSILKAKNAGMSLGAK 1081
Db 1128 GAAGPLPSEAVQWLC HQAFLKLTIRHRTVYVPLLSGLSRTAQTOLSRKLPCTTTALEAAA 1187
QY 1082 GAAGPLPSEAVQWLC HQAFLKLTIRHRTVYVPLLSGLSRTAQTOLSRKLPCTTTALEAAA 1141
Db 1188 NPALPSDFKTILD 1200
QY 1142 NPALPSDFKTILD 1154

RESULT 6
ID US-08-912-951-324 STANDARD; PRT; 1200 AA.
XX xxxxxx

Sequence 324, Application US/08912951

Sequence 324, Application US/08912951

GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC THERAPEUTIC METHODS
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017

CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-0026000US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 324:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1200 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1200 AA; 134322 MW; 7387257 CN;
Query Match 99.8%; Score 8607; DB 13; Length 1200;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 48 ASTQRCVLLRTWEALAPATPAMPRAPCRAVRSLSLSHREVLPLATFVRRRLGPGQWRLV 107
QY 2 ASGQRCVLLRTWEALAPATPAMPRAPCRAVRSLSLSHREVLPLATFVRRRLGPGQWRLV 61
Db 108 QRGDPAAFALVAQCLVCPWDARPPAPSPROVSCLELVARVLRQRCERGAKNVLA 167
QY 62 QRGDPAAFALVAQCLVCPWDARPPAPSPROVSCLELVARVLRQRCERGAKNVLA 121
Db 168 GFALLDGGARGGPEAFTTSVRSYLPNTVDALRGSGAWGLLRVDDVLLHLLARCALF 227
QY 122 GFALLDGGARGGPEAFTTSVRSYLPNTVDALRGSGAWGLLRVDDVLLHLLARCALF 181
Db 228 VLVAPSCAYQVCGPPLYQLGAATQARPPHAGSPRRRLGCERAWNHSVREAGVPLGLPAP 287
QY 182 VLVAPSCAYQVCGPPLYQLGAATQARPPHAGSPRRRLGCERAWNHSVREAGVPLGLPAP 241
Db 288 GARRGGASRSLLPLKPRRGAAPPEPTPVCGQSWAHFPGTRGSDRGFCVVSARPA 347
QY 242 GARRGGASRSLLPLKPRRGAAPPEPTPVCGQSWAHFPGTRGSDRGFCVVSARPA 301
Db 348 EEATSLLEGALSGTRHSHPSVGRHAGHPSTSRPPRPMDTPCPVYAEKHFLYSSGDKE 407
QY 302 EEATSLLEGALSGTRHSHPSVGRHAGHPSTSRPPRPMDTPCPVYAEKHFLYSSGDKE 361
Db 408 QLRPSFLLSLRPSLTGARRLVETIFLGSRPWMPGTPRRLPLRPLQRYQWMRPLFLELGN 467
QY 362 QLRPSFLLSLRPSLTGARRLVETIFLGSRPWMPGTPRRLPLRPLQRYQWMRPLFLELGN 421
Db 468 HAQCPGVLLKTHCPLRAAVTPAAGVCAREKPOGSAVAPEEEDTDPRRLVQLLRQHSPPW 527
QY 422 HAQCPGVLLKTHCPLRAAVTPAAGVCAREKPOGSAVAPEEEDTDPRRLVQLLRQHSPPW 481
Db 528 QVYGFVRACLRLVPPGLGSRHNRERLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 587
QY 482 QVYGFVRACLRLVPPGLGSRHNRERLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 541
Db 588 LRRSPGVCPVAAEHLREEILAKFLHLMMSVYVVELLSFFVYTTTFTQKNRLLFFYRKS 647
QY 542 LRRSPGVCPVAAEHLREEILAKFLHLMMSVYVVELLSFFVYTTTFTQKNRLLFFYRKS 601
Db 648 VWSKLSIGIRQHLKRVQLRELSEAEVROHREARPAALLTSRLRFPKPDGLRPVNDYV 707
QY 602 VWSKLSIGIRQHLKRVQLRELSEAEVROHREARPAALLTSRLRFPKPDGLRPVNDYV 661

Db 708 VGARTFRRKRAERLTSRVKALFSLVLYERARRPGLGASVLGLDDITHRAWRFTVLVRA 767
QY 662 VGARTFRRKRAERLTSRVKALFSLVLYERARRPGLGASVLGLDDITHRAWRFTVLVRA 721
Db 768 ODPPPELYFKVDVTGAYDIPQDRLTEVIASIKPONTYCVRRYAVVQAAHGHVRKAF 827
QY 722 ODPPPELYFKVDVTGAYDIPQDRLTEVIASIKPONTYCVRRYAVVQAAHGHVRKAF 781
Db 828 KSHVSTLTDLPYMRQFVAHLQETSPRLDVAWIEQSSSLEASGLDFVLFWMCHHAVR 887
QY 782 KSHVSTLTDLPYMRQFVAHLQETSPRLDVAWIEQSSSLEASGLDFVLFWMCHHAVR 841
Db 888 IRGKSYVQCGIPOGISTLLSLCYGDMENKLFAGIRRDGLLRVDDFLVTPHLTH 947
QY 842 IRGKSYVQCGIPOGISTLLSLCYGDMENKLFAGIRRDGLLRVDDFLVTPHLTH 901
Db 948 AKTFRLTLVRGVPYGCVMNLRKTVNFPVEDEALGTAFAVQMPAHGLFPWCGLLDTRT 1007
QY 902 AKTFRLTLVRGVPYGCVMNLRKTVNFPVEDEALGTAFAVQMPAHGLFPWCGLLDTRT 961
Db 1008 LEVQSDYSSYARTSIRASLTFNRGFKAGNRMRKLFGLVLRKCHSLFLDLQVNSLQTVCT 1067
QY 962 LEVQSDYSSYARTSIRASLTFNRGFKAGNRMRKLFGLVLRKCHSLFLDLQVNSLQTVCT 1021
Db 1068 NIKYKILLQAYRFHACVQLPFFHQVQWKNPTFFLRVISDTASLCYSILKAKNAGMSLGAK 1127
QY 1022 NIKYKILLQAYRFHACVQLPFFHQVQWKNPTFFLRVISDTASLCYSILKAKNAGMSLGAK 1081
Db 1128 GAAGPLPSEAVOMLCHOAFLLKLRHRTVYVPLIGSLRTAQTOLSRKLPGLTTTLEAAA 1187
QY 1082 GAAGPLPSEAVOMLCHOAFLLKLRHRTVYVPLIGSLRTAQTOLSRKLPGLTTTLEAAA 1141
Db 1188 NPALPSDFKTILD 1200
QY 1142 NPALPSDFKTILD 1154

RESULT 7
ID US-08-911-312-33 STANDARD; PRT: 1200 AA.
XX
AC xxxxxx
XX
DT
XX

Sequence 33, Application US/08911312

Sequence 33, Application US/08911312
GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William

TITLE OF INVENTION: Telomerase Reverse Transcriptase

NUMBER OF SEQUENCES: 170

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,312
FILING DATE: 14-AUG-1997

CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Einhorn, Gregory P.
CC REGISTRATION NUMBER: 38,440
CC REFERENCE/DOCKET NUMBER: 015389-002500US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 33:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1200 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1200 AA; 134322 MW; 7387257 CN;

Query Match 99.8%; Score 8607; DB 13; Length 1200;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 48 ASTQRCVLLRTWEALAPATPAMPAPRCRAVRSLRSHRYEVLPLATFVRRLLGPGQGRWLV 107
QY 2 ASQRCVLLRTWEALAPATPAMPAPRCRAVRSLRSHRYEVLPLATFVRRLLGPGQGRWLV 61
Db 108 QRGDPAAFRALVAQCLVCPWDARPPAPSPRQVSCLELVARVQLRCERGAKNVLA 167
QY 62 QRGDPAAFRALVAQCLVCPWDARPPAPSPRQVSCLELVARVQLRCERGAKNVLA 121
Db 168 GFALLDGARGGPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGGDDVLVHLLARCALF 227
QY 122 GFALLDGARGGPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGGDDVLVHLLARCALF 181
Db 228 VLVAPSCAYQVCGPPLYQLGAATQARPPHASPGRRLRCERAWNHSVREAGVPLGLPAP 287
QY 182 VLVAPSCAYQVCGPPLYQLGAATQARPPHASPGRRLRCERAWNHSVREAGVPLGLPAP 241
Db 288 GARRRGGSASRSILPLKPRPRGAAPPEPTPVQSGSWAHPGTRGSDRGFCVSPARPA 347
QY 242 GARRRGGSASRSILPLKPRPRGAAPPEPTPVQSGSWAHPGTRGSDRGFCVSPARPA 301
Db 348 EBATSLGALSGTRHSHPSVGRHAGPSTSRPPPPMDTPCPVYAEKHFLLYSSGDKE 407
QY 302 EBATSLGALSGTRHSHPSVGRHAGPSTSRPPPPMDTPCPVYAEKHFLLYSSGDKE 361
Db 408 QLRPSFLLSRLPSLTGARRLVETIFLGSRPWNMPGTTPRLPLPQRYWQMRPLFLELGN 467
QY 362 QLRPSFLLSRLPSLTGARRLVETIFLGSRPWNMPGTTPRLPLPQRYWQMRPLFLELGN 421
Db 468 HAQCPYGVLLKTHCPPLRAAVTPAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSSPW 527
QY 422 HAQCPYGVLLKTHCPPLRAAVTPAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSSPW 481

Db 528 QVYGFVACLRRLVPPGLMGSRHNERFLRNTKFKISLGHAKLSLOELTWKMSVRDCAW 587
QY 482 QVYGFVACLRRLVPPGLMGSRHNERFLRNTKFKISLGHAKLSLOELTWKMSVRDCAW 541
Db 588 LRSPGVGCVPAAEHRLREILAKFLHMLSVVVELLSRFFVYVTEFTFKNRLEFYRKS 647
QY 542 LRSPGVGCVPAAEHRLREILAKFLHMLSVVVELLSRFFVYVTEFTFKNRLEFYRKS 601
Db 648 VWSKLSIGIRQLKRVQLRELSEAEVQRHREARPAALLTSRLRFIPKPDGLRPIVNMYY 707
QY 602 VWSKLSIGIRQLKRVQLRELSEAEVQRHREARPAALLTSRLRFIPKPDGLRPIVNMYY 661
Db 708 VGARTFRERKRAELTSRVKALFSLVLYNERARRPGLLGASVLGDDIHRARWTFVLVR 767
QY 662 VGARTFRERKRAELTSRVKALFSLVLYNERARRPGLLGASVLGDDIHRARWTFVLVR 721
Db 768 QDPPELYFVKVDVTGAYDIPDRLTEVIASIKPONTYCVRRYAVVQAAHGHVKAF 827
QY 722 QDPPELYFVKVDVTGAYDIPDRLTEVIASIKPONTYCVRRYAVVQAAHGHVKAF 781
Db 828 KSHVSTLTDLPYMRQFVAHLQETSPLRDVAVIEQSSSLNEASSGLDFVFLRFMCHHAVR 887
QY 782 KSHVSTLTDLPYMRQFVAHLQETSPLRDVAVIEQSSSLNEASSGLDFVFLRFMCHHAVR 841
Db 888 IRKSVYQCOGIPQGSILSTLLSLCYGDMENKLFAGIRRDGLLLRLVDDFLVTPHLTH 947
QY 842 IRKSVYQCOGIPQGSILSTLLSLCYGDMENKLFAGIRRDGLLLRLVDDFLVTPHLTH 901
Db 948 AKTFELTLVRGVPEYGCNNLRKTVNFPVEDEALGTAFTVQMPAHGLFPWCGLLDTRT 1007
QY 902 AKTFELTLVRGVPEYGCNNLRKTVNFPVEDEALGTAFTVQMPAHGLFPWCGLLDTRT 961
Db 1008 LEVQSDYSSVARTSIRASLTFRNGFRAGRNMRKLFGLVRLKCHSLFLDLQVNSLOTVCT 1067
QY 962 LEVQSDYSSVARTSIRASLTFRNGFRAGRNMRKLFGLVRLKCHSLFLDLQVNSLOTVCT 1021
Db 1068 NIKYKILLQAYRHACVQLQPFHQVWKNPTFFELRVISDTSALCYSLTKAKNAGMSLGAK 1127
QY 1022 NIKYKILLQAYRHACVQLQPFHQVWKNPTFFELRVISDTSALCYSLTKAKNAGMSLGAK 1081
Db 1128 GAAGPLPSEAVOMLCHOAFLKLTIRHRTVYVPLGLSLRTAQTOLSRKLPCTTITALEAAA 1187
QY 1082 GAAGPLPSEAVOMLCHOAFLKLTIRHRTVYVPLGLSLRTAQTOLSRKLPCTTITALEAAA 1141
Db 1188 NPALPSDFKTILD 1200
QY 1142 NPALPSDFKTILD 1154

RESULT 8
ID US-08-974-549-600 STANDARD; PRT; 1285 AA.

AC xxxxxx

DT

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 600, Application US/08974549

Sequence 600, Application US/08974549

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin B.

APPLICANT: Andrews, William H.

TITLE OF INVENTION: Human Telomerase Catalytic Subunit

NUMBER OF SEQUENCES: 726

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent in Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/974,549
CC FILING DATE: 19-NOV-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph Ted
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-0026100S
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 600:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1285 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: Protein
CC LOCATION: 1..1285
CC OTHER INFORMATION: /note= "fusion protein composed of
CC OTHER INFORMATION: enterokinase cleavable, His tagged
CC OTHER INFORMATION: thioredoxin moiety and full length hTERT"
CC SEQUENCE 1285 AA; 143529 MW; 849280 CN;

Query Match 99.8%; Score 8607; DB 18; Length 1285;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 133 ASTQRCVLLTWEALAPATPAMPAPRCRAVRSLLSHRYEVLPLATFVRRLGPGQGRV 192

QY 2 ASQRCVLLTWEALAPATPAMPAPRCRAVRSLLSHRYEVLPLATFVRRLGPGQGRV 61

| | | | |
|----|------|---|------|
| Db | 193 | QRGDPAAFRALVAQCLVCPWMDARPPAAPSPQVSCLUKELVARVQLRJCERGAKNVLAF | 252 |
| Qy | 62 | QRGDPAAFRALVAQCLVCPWMDARPPAAPSPQVSCLUKELVARVQLRJCERGAKNVLAF | 121 |
| Db | 253 | GFALLDGCARGGPPPEAFTTSVRSYLPNTVTDALRGSGANGLLLRRVGDVVLHLLARCALF | 312 |
| Qy | 122 | GFALLDGCARGGPPPEAFTTSVRSYLPNTVTDALRGSGANGLLLRRVGDVVLHLLARCALF | 181 |
| Db | 313 | VLVAPSCAYQVCPPPLYQLGAATQARPPPHASGPRRRRLCERAMNHSVREAGVPLGLPAP | 372 |
| Qy | 182 | VLVAPSCAYQVCPPPLYQLGAATQARPPPHASGPRRRRLCERAMNHSVREAGVPLGLPAP | 241 |
| Db | 373 | GARRGGSGRSLLPLPKRPRRGAAPPERTPVQGSWAHPGRTGRGSCVVSARPAP | 432 |
| Qy | 242 | GARRGGSGRSLLPLPKRPRRGAAPPERTPVQGSWAHPGRTGRGSCVVSARPAP | 301 |
| Db | 433 | EEATSLLEGALSGTRHSHPSVGRQHHAGPSTSRPPRPMDTCPVPYATKHFLLSSGDKE | 492 |
| Qy | 302 | EEATSLLEGALSGTRHSHPSVGRQHHAGPSTSRPPRPMDTCPVPYATKHFLLSSGDKE | 361 |
| Db | 493 | QLRPSFLLSLRSLTGARRLVETIFLGSRPWMPGTPRRLPLRPORYWQMRPLFLELLGN | 552 |
| Qy | 362 | QLRPSFLLSLRSLTGARRLVETIFLGSRPWMPGTPRRLPLRPORYWQMRPLFLELLGN | 421 |
| Db | 553 | HAQCPYGVLLKTHCPRLAAVTPAAGYACAREKPGQSVAAPEEEDTPRRLVLQLRHSSPW | 612 |
| Qy | 422 | HAQCPYGVLLKTHCPRLAAVTPAAGYACAREKPGQSVAAPEEEDTPRRLVLQLRHSSPW | 481 |
| Db | 613 | QVYGFVACLRLLVPPGLWGSRRNERRFLRNTKKFISLGKHAHKLSQLBTWKMSVROCAW | 672 |
| Qy | 482 | QVYGFVACLRLLVPPGLWGSRRNERRFLRNTKKFISLGKHAHKLSQLBTWKMSVROCAW | 541 |
| Db | 673 | LRRSPGVCVPAAEHRLREELIAKFLHLLMSVYVELLRSFPYVTTETTFQKNRLLFFYRKS | 732 |
| Qy | 542 | LRRSPGVCVPAAEHRLREELIAKFLHLLMSVYVELLRSFPYVTTETTFQKNRLLFFYRKS | 601 |
| Db | 733 | VWSKLQSIGIRQHLKRVQURELSEAEVRRQREARPAALLTSRLRFIPKPDGLRPVNMDDY | 792 |
| Qy | 602 | VWSKLQSIGIRQHLKRVQURELSEAEVRRQREARPAALLTSRLRFIPKPDGLRPVNMDDY | 661 |
| Db | 793 | VGARTFRREKRAERLSRVKALFESVLYNEARRRPGLLGASVLGLDDIIRAWRTFVLRVRA | 852 |
| Qy | 662 | VGARTFRREKRAERLSRVKALFESVLYNEARRRPGLLGASVLGLDDIIRAWRTFVLRVRA | 721 |
| Db | 853 | QDPPPELYFVKVDVTGAYDTIPODRUTEVIAISIKPONTYCVRRYAVVQKAAGHVRKAF | 912 |
| Qy | 722 | QDPPPELYFVKVDVTGAYDTIPODRUTEVIAISIKPONTYCVRRYAVVQKAAGHVRKAF | 781 |
| Db | 913 | KSHVSTLTLQPMYMRQFVAHLQETSPLRDAVYIEQSSSLNEASSGLDFVFLRFMCHHAVR | 972 |
| Qy | 782 | KSHVSTLTLQPMYMRQFVAHLQETSPLRDAVYIEQSSSLNEASSGLDFVFLRFMCHHAVR | 841 |
| Db | 973 | IRGKSYVQOCIGIPQGSITLLCSLCYGMENKFLFAGIRRGILLURLVDDFLLVTPHLTH | 1032 |
| Qy | 842 | IRGKSYVQOCIGIPQGSITLLCSLCYGMENKFLFAGIRRGILLURLVDDFLLVTPHLTH | 901 |
| Db | 1033 | AKTFLRLTVRGVPEYGCVNLRKTVYVPEDEALGGTAFVQMPAHGLFPWCGLLLDTRT | 1092 |
| Qy | 902 | AKTFLRLTVRGVPEYGCVNLRKTVYVPEDEALGGTAFVQMPAHGLFPWCGLLLDTRT | 961 |
| Db | 1093 | LEVQSDYSYARTSIRASLTFNKGFKAGNRRMRKFLGVLRKCHSLFLDLQVNSLQTVCT | 1152 |
| Qy | 962 | LEVQSDYSYARTSIRASLTFNKGFKAGNRRMRKFLGVLRKCHSLFLDLQVNSLQTVCT | 1021 |
| Db | 1153 | NIYKILLQAYRFHACVQLPFFHQQVWKNPTEFLRVISDTASLCYSILKAKNAGMSLGAK | 1212 |
| Qy | 1022 | NIYKILLQAYRFHACVQLPFFHQQVWKNPTEFLRVISDTASLCYSILKAKNAGMSLGAK | 1081 |
| Db | 1213 | GAAGPLPSEAVOMLCHQAFLLKLTRHRVTYVPLGLSLRTAOPOLSRKPLPGTTLTALFAAA | 1272 |
| Qy | 1082 | GAAGPLPSEAVOMLCHQAFLLKLTRHRVTYVPLGLSLRTAOPOLSRKPLPGTTLTALFAAA | 1141 |
| Db | 1273 | NPALPSDFKTIILD | 1285 |

CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 1285 AA: 143529 MW: 8449280 CN;

Query Match 99.8%; Score 8607; DB 13; Length 1285;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 133 ASTQRCVLLTWTALPATAMPAPRCRAVRSLLRSHYREVLPPLATFVRRLGPGQWRV 192
|||
QY 2 ASQRCVLLTWTALPATAMPAPRCRAVRSLLRSHYREVLPPLATFVRRLGPGQWRV 61
|||
Db 193 QRGDPAFRAVLAQCLVCVPWDARPPAPSFQVSCCLKELVARVLRQRCERGAKNVLA 252
|||
QY 62 QRGDPAFRAVLAQCLVCVPWDARPPAPSFQVSCCLKELVARVLRQRCERGAKNVLA 121
|||
Db 253 GFALLDGARGPPEAFETTSVRSYLPNTVTDALRGSGAWGLLLRRVGDVVLVHLLARCALF 312
|||
QY 122 GFALLDGARGPPEAFETTSVRSYLPNTVTDALRGSGAWGLLLRRVGDVVLVHLLARCALF 181
|||
Db 313 VLVAPSCAYQVCGPPLYLQGAATQARPPPHAGSPRRRLGCRERAWNHSVREAGVPLGLPAP 372
|||
QY 182 VLVAPSCAYQVCGPPLYLQGAATQARPPPHAGSPRRRLGCRERAWNHSVREAGVPLGLPAP 241
|||
Db 373 GARRGGASRSRLPLKRRPRRGAPEPTPVGGSWAHGPRTRGSDRGFCVVSARPA 432
|||
QY 242 GARRGGASRSRLPLKRRPRRGAPEPTPVGGSWAHGPRTRGSDRGFCVVSARPA 301
|||
Db 433 EEATSLGALSGTRHSHPSVGRQHAGPSTSRPPRPWDTPCPVVAETHFLYSSGDKE 492
|||
QY 302 EEATSLGALSGTRHSHPSVGRQHAGPSTSRPPRPWDTPCPVVAETHFLYSSGDKE 361
|||
Db 493 QLRSFLLSLRLSLTGARLVETIFIGSRPMPGTPRRLPRLPQRYQWMPRLFLELLGN 552
|||
QY 362 QLRSFLLSLRLSLTGARLVETIFIGSRPMPGTPRRLPRLPQRYQWMPRLFLELLGN 421
|||
Db 553 HAQCPYGLKTKCHPLRAAYTPAAGVCAREKPGQSVAAPEEEDTPRLVOLLRQHSSPW 612
|||
QY 422 HAQCPYGLKTKCHPLRAAYTPAAGVCAREKPGQSVAAPEEEDTPRLVOLLRQHSSPW 481
|||
Db 613 QVTGFVRACLRLVPPGLGWSRHNREFLNTRKFFISLGHAKLSLQELTWKMSVRDCAW 672
|||
QY 482 QVTGFVRACLRLVPPGLGWSRHNREFLNTRKFFISLGHAKLSLQELTWKMSVRDCAW 541
|||
Db 673 LRSPGVCVPAAEHRLREILAKFLHLMVSVYVELLSFFYYTTFQKNRLLFFYRKS 732
|||
QY 542 LRSPGVCVPAAEHRLREILAKFLHLMVSVYVELLSFFYYTTFQKNRLLFFYRKS 601
|||
Db 733 VWSKLSIGIRQHLKRVQLRELSEAEVQROHREAPALLTSRLRFIPKPDGLRPIVNMVY 792
|||
QY 602 VWSKLSIGIRQHLKRVQLRELSEAEVQROHREAPALLTSRLRFIPKPDGLRPIVNMVY 661
|||
Db 793 VGARTFREKRAERLTSRKALFSLVLYERARRPGLLGASVLGLDDIHRARWTFVLVRRA 852
|||
QY 662 VGARTFREKRAERLTSRKALFSLVLYERARRPGLLGASVLGLDDIHRARWTFVLVRRA 721
|||
Db 853 QDPPPELYFVKVDVTGAYDTIPQDLTEVIASIKPQNTYCVRRYAVYQKAAGHVRKAF 912
|||
QY 722 QDPPPELYFVKVDVTGAYDTIPQDLTEVIASIKPQNTYCVRRYAVYQKAAGHVRKAF 781
|||
Db 913 KSHVSTLTDLPQYNRQFVAHLQETSPRDVAVYQSSSLNEASGLDFVLRPMCHAVR 972
|||
QY 782 KSHVSTLTDLPQYNRQFVAHLQETSPRDVAVYQSSSLNEASGLDFVLRPMCHAVR 841
|||
Db 973 IRGKSYVOCOGIPOGSIILSLCLSGYGDMEKLFAGIRRDGLLLRLVDDFLVTPPLTH 1032
|||
QY 842 IRGKSYVOCOGIPOGSIILSLCLSGYGDMEKLFAGIRRDGLLLRLVDDFLVTPPLTH 901
|||
Db 1033 AKTFRLTLVRGPEYGCWNLRTKTVVNFVEDEALGCTAFVQMPAHGLFPWCGLLDTRT 1092
|||
QY 902 AKTFRLTLVRGPEYGCWNLRTKTVVNFVEDEALGCTAFVQMPAHGLFPWCGLLDTRT 961
|||
Db 1093 LEVQSDYSSYARTSIRASLTFRNGFKAGRNMRKLFVGLRLKCHSLFLDLQVNSLQTVCT 1152
|||

QY 962 LEVQSDYSSYARTSIRASLTFRNGFKAGRNMRKLFVGLRLKCHSLFLDLQVNSLQTVCT 1021
|||
Db 1153 NIYKILLQAYRFHACVQLQPFHQQVWKNPFTFFRLRVISDTASLCYSILKAKNAGMSLGAK 1212
|||
QY 1022 NIYKILLQAYRFHACVQLQPFHQQVWKNPFTFFRLRVISDTASLCYSILKAKNAGMSLGAK 1081
|||
Db 1213 GAAGPLPSEAVOWLCHQAFLLKLTNRHRTVYVPLLSRLTAQTLRSKLPGLTTLTALEAAA 1272
|||
QY 1082 GAAGPLPSEAVOWLCHQAFLLKLTNRHRTVYVPLLSRLTAQTLRSKLPGLTTLTALEAAA 1141
|||
Db 1273 NPALPSDFKTILD 1285
|||
QY 1142 NPALPSDFKTILD 1154
|||

RESULT 10
ID US-08-912-951-314 STANDARD; PRT; 1285 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 314, Application US/08912951
XX
CC Sequence 314, Application US/08912951
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC TITL OF INVENTION: THERAPEUTIC METHODS
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002600US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 314:
CC .SEQUENCE CHARACTERISTICS:
CC LENGTH: 1285 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SQ SEQUENCE 1285 AA; 143529 MW; 8449280 CN;

Query Match 99.8%; Score 8607; DB 13; Length 1285;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 133 ASTORCVLLRTWEALAPATPAMPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRVY 192
|||
QY 2 ASGQRCVLLRTWEALAPATPAMPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRVY 61
|||
Db 193 QRGDPAAFRALVAQCLVCPVMDARPPAPAFSFRQVSCIKELVAVLQRLCERGAKNVLA 252
|||
QY 62 QRGDPAAFRALVAQCLVCPVMDARPPAPAFSFRQVSCIKELVAVLQRLCERGAKNVLA 121
|||
Db 253 GFALLDGARGPPPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGDVLLVHLLARCALF 312
|||
QY 122 GFALLDGARGPPPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGDVLLVHLLARCALF 181
|||
Db 313 VLVAPSCAYQVCGPPLVQLGAATQARPPPHAGSPRRRLGGERAWNHVSREAGVPLGLPAP 372
|||
QY 182 VLVAPSCAYQVCGPPLVQLGAATQARPPPHAGSPRRRLGGERAWNHVSREAGVPLGLPAP 241
|||
Db 373 GARRGGSASRLPLKRPREGAAPERPTPVGOGSWAHGPGRTGPDRCFCVVSARPA 432
|||
QY 242 GARRGGSASRLPLKRPREGAAPERPTPVGOGSWAHGPGRTGPDRCFCVVSARPA 301
|||
Db 433 EEATSLGALSGRHSPSVGROHAGPPTSRPPRPWDTPCPVYAETKHFLYSSGDKE 492
|||
QY 302 EEATSLGALSGRHSPSVGROHAGPPTSRPPRPWDTPCPVYAETKHFLYSSGDKE 361
|||
Db 493 QLRSFLLSLRPSLTGARLVETIFIGSRPMPGTPRRLPRLPQRYWQMRPLFLELLGN 552
|||
QY 362 QLRSFLLSLRPSLTGARLVETIFIGSRPMPGTPRRLPRLPQRYWQMRPLFLELLGN 421
|||
Db 553 HAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPGQSVAAPEEEDTPRLVOLLRQHSSPW 612
|||
QY 422 HAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPGQSVAAPEEEDTPRLVOLLRQHSSPW 481
|||
Db 613 QVYGFVRACLRLVPPGLWGSRRHNERFLNRTKFIISLGHAKLSLQELTWKMSVRDCAW 672
|||
QY 482 QVYGFVRACLRLVPPGLWGSRRHNERFLNRTKFIISLGHAKLSLQELTWKMSVRDCAW 541
|||
Db 673 LRRSPGVCVPAAEHRLREELAKFLHLMVSVYVVELLSFFYVTTFTQKNRFFYRKS 732
|||
QY 542 LRRSPGVCVPAAEHRLREELAKFLHLMVSVYVVELLSFFYVTTFTQKNRFFYRKS 601
|||
Db 733 VWSKLQSIGIRHLLKRYQLRELSAEVROHREARPAALLTSRLRFPKPDGLRPTVNDYV 792
|||
QY 602 VWSKLQSIGIRHLLKRYQLRELSAEVROHREARPAALLTSRLRFPKPDGLRPTVNDYV 661
|||
Db 793 VGARTFRREKRAELTSRVKALFSLVLYNERARRPGLLGASVGLGDDTHRAWRFLVLRVA 852
|||
QY 662 VGARTFRREKRAELTSRVKALFSLVLYNERARRPGLLGASVGLGDDTHRAWRFLVLRVA 721
|||
Db 853 QDPPPELYFVKVDVGTAYDITIPQDLREVIASIIKQNTYCVRRYAVVQAAHGHVRKAF 912
|||
QY 722 QDPPPELYFVKVDVGTAYDITIPQDLREVIASIIKQNTYCVRRYAVVQAAHGHVRKAF 781
|||
Db 913 KSHVSTLTDLPQYMRQFVAHLQETSPLRDAVVVIEQSSSLEASSGLDFVLRFMCHHAVR 972
|||

QY 782 KSHVSTLTDLPQYMRQFVAHLQETSPLRDAVVVIEQSSSLEASSGLDFVLRFMCHHAVR 841
|||
Db 973 IRGKSYVQCGIPIQSGSILSTLLCSLCYGDWENKLFAGIRRDGILLRLVDVDFLLVTPHLTH 1032
|||
QY 842 IRGKSYVQCGIPIQSGSILSTLLCSLCYGDWENKLFAGIRRDGILLRLVDVDFLLVTPHLTH 901
|||
Db 1033 AKTFELTLVRGPEYGCYVNLKRTVVNFVDEALGGTAFVQMPAHGLFPWCGLLDTRT 1092
|||
QY 902 AKTFELTLVRGPEYGCYVNLKRTVVNFVDEALGGTAFVQMPAHGLFPWCGLLDTRT 961
|||
Db 1093 LEVQSDYSYVARTSIRASLTFRNGFKAGNMRKLFGLVRLKCHSLFLDQVNSLQTVCT 1152
|||
QY 962 LEVQSDYSYVARTSIRASLTFRNGFKAGNMRKLFGLVRLKCHSLFLDQVNSLQTVCT 1021
|||
Db 1153 NIYKILLQAYRFHACVQLQPFHQQVKNPTFFLRVISTASLCYSILKAKNAGMSLGAK 1212
|||
QY 1022 NIYKILLQAYRFHACVQLQPFHQQVKNPTFFLRVISTASLCYSILKAKNAGMSLGAK 1081
|||
Db 1213 GAAGPLPSEAVOWMLCHQAFLLKLTLRHRTVYVPLLSGLRTAQOLSKRLPGTTLTALAAAA 1272
|||
QY 1082 GAAGPLPSEAVOWMLCHQAFLLKLTLRHRTVYVPLLSGLRTAQOLSKRLPGTTLTALAAAA 1141
|||
Db 1273 NPALPSDFKTILD 1285
|||
QY 1142 NPALPSDFKTILD 1154
|||

RESULT 11
ID US-08-911-312-55 STANDARD; PRT; 1407 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 55, Application US/08911312
XX
XX Sequence 55, Application US/08911312
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase
CC NUMBER OF SEQUENCES: 170
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/911,312
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997

CC APPLICATION NUMBER: US/08/974,549
CC FILING DATE: 19-NOV-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 08-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph Ted
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002610US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 628:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1407 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: Protein
CC LOCATION: 1..1407
CC OTHER INFORMATION: /note= "fusion protein composed of
CC OTHER INFORMATION: enhanced green fluorescent protein
CC OTHER INFORMATION:
CC OTHER INFORMATION: 5' untranslated region of hTERT mRNA and
CC OTHER INFORMATION: hTERT protein sequence"
CC SEQUENCE 1407 AA; 157668 MW; 10134798 CN;

Query Match 99.6%; Score 8590; DB 18; Length 1407;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 1150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 255 ASTORCVLLTWTWEALAPATAMPAPRCRAVRSLLRSHYREVLPATFVRRLPGQGRVLY 314
QY 2 ASGRCVLLTWTWEALAPATAMPAPRCRAVRSLLRSHYREVLPATFVRRLPGQGRVLY 61
Db 315 QRGDPAAFRALVAQVLCVCPWDARPPAAPSFROVSCCLKELVARVLQRCERGAKNVLA 374
QY 62 QRGDPAAFRALVAQVLCVCPWDARPPAAPSFROVSCCLKELVARVLQRCERGAKNVLA 121
Db 375 GFALLDARGGPPPEAFTTSVRSYLPNTVTDALRGSGAWGALLRRVGDVVLVHLLARCALF 434
QY 122 GFALLDARGGPPPEAFTTSVRSYLPNTVTDALRGSGAWGALLRRVGDVVLVHLLARCALF 181

Db 435 VLVAPSCAYOVCGPPLYQLGAATQARPPPHASGPRRRRLCCERAWNHSVREAGVPLGLPAP 494
QY 182 VLVAPSCAYOVCGPPLYQLGAATQARPPPHASGPRRRRLCCERAWNHSVREAGVPLGLPAP 241
Db 495 GARRRGGSASRSLPLPKRPRRGAAPERTPVGQSWAHPTGRGSDRGFCFVVSAPARA 554
QY 242 GARRRGGSASRSLPLPKRPRRGAAPERTPVGQSWAHPTGRGSDRGFCFVVSAPARA 301
Db 555 EBATSLEGALSGTRHSHPSVGRQHAGPPSTSRPPRPMDTPCPVPVYAEKHFLYSSGDKE 614
QY 302 EBATSLEGALSGTRHSHPSVGRQHAGPPSTSRPPRPMDTPCPVPVYAEKHFLYSSGDKE 361
Db 615 OLRPSELLSLRPSLTGARRLVETIFLGSRPMPGTPRRLPRLPQRYWQMRPLFLELLGN 674
QY 362 OLRPSELLSLRPSLTGARRLVETIFLGSRPMPGTPRRLPRLPQRYWQMRPLFLELLGN 421
Db 675 HAQCPYGVLLKTHCPLRAAATPAAGVCAREKPGQSVAAPEEEDTDPRRLVQLLRHSSPW 734
QY 422 HAQCPYGVLLKTHCPLRAAATPAAGVCAREKPGQSVAAPEEEDTDPRRLVQLLRHSSPW 481
Db 735 QYGVFVRACLRLVPPGLWGSRHNERFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 794
QY 482 QYGVFVRACLRLVPPGLWGSRHNERFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 541
Db 795 LRSPGCVGCPAAEHRLREEILAKELHMLMSVYVVELLSRFFYVTTTQKNRLEFFYRS 854
QY 542 LRSPGCVGCPAAEHRLREEILAKELHMLMSVYVVELLSRFFYVTTTQKNRLEFFYRS 601
Db 855 VMSKLSIGIRQHLKRVQLRELSEAVRQREARPAALTSRLRFIPKPDGLRPIVNDYV 914
QY 602 VMSKLSIGIRQHLKRVQLRELSEAVRQREARPAALTSRLRFIPKPDGLRPIVNDYV 661
Db 915 VGARTFRREKRAERLTSRVKALFSVLNLERARRPGLLGASVLGDDIHRARWTFVLVRRA 974
QY 662 VGARTFRREKRAERLTSRVKALFSVLNLERARRPGLLGASVLGDDIHRARWTFVLVRRA 721
Db 975 QDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIIPQNTYCVRRYAVVQAAHGHVKAF 1034
QY 722 QDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIIPQNTYCVRRYAVVQAAHGHVKAF 781
Db 1035 KSHVSTLTLQPYMRQFVAHLQETSPDRDAVVEQSSSNEASSGLFDVFLRFMCHHAVR 1094
QY 782 KSHVSTLTLQPYMRQFVAHLQETSPDRDAVVEQSSSNEASSGLFDVFLRFMCHHAVR 841
Db 1095 IRKSYVQCGIPIQGSILSTLCSLCYGDENKLFAGIRRDGLLRVLDVDFLLVTPHLTH 1154
QY 842 IRKSYVQCGIPIQGSILSTLCSLCYGDENKLFAGIRRDGLLRVLDVDFLLVTPHLTH 901
Db 1155 AKTFLETLVRGYPEYGCYVNLKTKTVNPFVEDEALGGTAFVQMPAHGLFPWCGLLLDTRT 1214
QY 902 AKTFLETLVRGYPEYGCYVNLKTKTVNPFVEDEALGGTAFVQMPAHGLFPWCGLLLDTRT 961
Db 1215 LEVQSDYSSYARTSIRASVTFNKRGFKAGRMRRKLFGLVLRKCHSLFLDLQVNSLQTVC 1274
QY 962 LEVQSDYSSYARTSIRASVTFNKRGFKAGRMRRKLFGLVLRKCHSLFLDLQVNSLQTVC 1021
Db 1275 NIYKILLQAYRFHACVQLQPFHQYQWKNPFTFFLRVVISDTASLCYSILKAKNAGMSLGAK 1334
QY 1022 NIYKILLQAYRFHACVQLQPFHQYQWKNPFTFFLRVVISDTASLCYSILKAKNAGMSLGAK 1081
Db 1335 GAAGPLPSEAVQWLCHQAFLLKTLHRVTYVPLLSLRTAQQLSRKLPGLTTLTALEAAA 1394
QY 1082 GAAGPLPSEAVQWLCHQAFLLKTLHRVTYVPLLSLRTAQQLSRKLPGLTTLTALEAAA 1141
Db 1395 NPALPSDFKTTILD 1407
QY 1142 NPALPSDFKTTILD 1154

RESULT 13
ID US-08-912-951-334 STANDARD; PRT; 1407 AA.
XX
AC xxxxxx

xx Sequence 334, Application US/08912951
dt Sequence 334, Application US/08912951
xx GENERAL INFORMATION:
cc APPLICANT: Cech, Thomas R.
cc APPLICANT: Lingner, Joachim
cc APPLICANT: Nakamura, Toru
cc APPLICANT: Chapman, Karen B.
cc APPLICANT: Morin, Gregg B.
cc APPLICANT: Harley, Calvin
cc APPLICANT: Andrews, William H.
cc TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
cc TITLE OF INVENTION: THERAPEUTIC METHODS
cc NUMBER OF SEQUENCES: 335
cc CORRESPONDENCE ADDRESS:
cc ADDRESSEE: Townsend and Crew LLP
cc STREET: Two Embarcadero Center, 8th Floor
cc CITY: San Francisco
cc STATE: California
cc COUNTRY: United States of America
cc ZIP: 94111
cc COMPUTER READABLE FORM:
cc MEDIUM TYPE: Floppy disk
cc COMPUTER: IBM PC compatible
cc OPERATING SYSTEM: PC-DOS/MS-DOS
cc SOFTWARE: PatentIn Release #1.0, Version #1.30
cc CURRENT APPLICATION DATA:
cc APPLICATION NUMBER: US/08/912,951
cc FILING DATE: 14-AUG-1997
cc CLASSIFICATION: 435
cc PRIOR APPLICATION DATA:
cc APPLICATION NUMBER: US 08/854,050
cc FILING DATE: 09-MAY-1997
cc CLASSIFICATION: 435
cc PRIOR APPLICATION DATA:
cc APPLICATION NUMBER: US 08/851,843
cc FILING DATE: 06-MAY-1997
cc CLASSIFICATION: 435
cc PRIOR APPLICATION DATA:
cc APPLICATION NUMBER: US 08/846,017
cc FILING DATE: 25-APR-1997
cc CLASSIFICATION: 435
cc PRIOR APPLICATION DATA:
cc APPLICATION NUMBER: US 08/844,419
cc FILING DATE: 18-APR-1997
cc CLASSIFICATION: 435
cc PRIOR APPLICATION DATA:
cc APPLICATION NUMBER: US 08/724,643
cc FILING DATE: 01-OCT-1996
cc CLASSIFICATION: 435
cc ATTORNEY/AGENT INFORMATION:
cc NAME: Apple, Randolph T.
cc REGISTRATION NUMBER: 36,429
cc REFERENCE/DOCKET NUMBER: 015389-002600US
cc TELECOMMUNICATION INFORMATION:
cc TELEPHONE: (415) 576-0200
cc TELEFAX: (415) 576-0300
cc INFORMATION FOR SEQ ID NO: 334:
cc SEQUENCE CHARACTERISTICS:
cc LENGTH: 1407 amino acids
cc TYPE: amino acid
cc STRANDEDNESS:
cc TOPOLOGY: linear
cc MOLECULE TYPE: protein
cc SEQUENCE 1407 AA; 157668 MW; 10134798 CN;
Query Match 99.6%; Score 8590; DB 13; Length 1407;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 1150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

|||||
QY 1082 GAAGLPSEAVQWLCQAFLLKLFHRVYVYPLGLSLRTAQTLQSRKLPGLTTLALEAAA 1141
Db 1395 NPALPSDFKTLID 1407
|||||
QY 1142 NPALPSDFKTLID 1154
RESULT 14
ID US-08-974-549-2 STANDARD; PRT; 1132 AA.
XX
AC
AC
XX
XX
DT
XX
XX
Sequence 2, Application US/08974549
XX
XX
Sequence 2, Application US/08974549
CC
CC
GENERAL INFORMATION:
CC
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Human Telomerase Catalytic Subunit
CC NUMBER OF SEQUENCES: 726
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/974,549
CC FILING DATE: 19-NOV-1997
CC CLASSIFICATION: 536
CC
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885

CC
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph Ted
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002610US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1132 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 1132 AA; 126995 MW; 6588209 CN;

Query Match 98.2%; Score 8465; DB 18; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MPRAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRLVQRGDPAAFRALVAOCLVCPW 60
|||||
QY 23 MPRAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRLVQRGDPAAFRALVAOCLVCPW 82
|||||
Db 61 DARPPPAAPSFROVSCLELVARVLQRLCERGAKNVLAFGFALLDARGGPPPEAFTTSVR 120
|||||
QY 83 DARPPPAAPSFROVSCLELVARVLQRLCERGAKNVLAFGFALLDARGGPPPEAFTTSVR 142
|||||
Db 121 SYLPNTVTDALRGSGAWGLLRVGGDVLVHLLARCALFVLVAPSCAYOVCGPPLVQLGA 180
|||||
QY 143 SYLPNTVTDALRGSGAWGLLRVGGDVLVHLLARCALFVLVAPSCAYOVCGPPLVQLGA 202
|||||
Db 181 ATQARPPPHASGPRRLGGERAWNHSVREAGVPLGLPAPGARRRGSGASRSPLPKRPRR 240
|||||
QY 203 ATQARPPPHASGPRRLGGERAWNHSVREAGVPLGLPAPGARRRGSGASRSPLPKRPRR 262
|||||
Db 241 GAAPERTPVGGSWAHFGRTRGSDRGFCVVSVPARPAEATSLGALSGRHSHPSVG 300
|||||
QY 263 GAAPERTPVGGSWAHFGRTRGSDRGFCVVSVPARPAEATSLGALSGRHSHPSVG 322
|||||
Db 301 ROHAGPPSTSRPPRWDTPCPVYAEKHFYSSGDKQLRPSFLLSLRPSLTGARRL 360
|||||
QY 323 ROHAGPPSTSRPPRWDTPCPVYAEKHFYSSGDKQLRPSFLLSLRPSLTGARRL 382
|||||
Db 361 VETIFLGSRPMPGTPRRLPRLPQRYWQMRPLELLELGHNAQCPYGVLLKTHCPRAAVT 420
|||||
QY 383 VETIFLGSRPMPGTPRRLPRLPQRYWQMRPLELLELGHNAQCPYGVLLKTHCPRAAVT 442
|||||
Db 421 PAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSPPQVYGFVACLRRLVPPGLWGS 480
|||||
QY 443 PAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSPPQVYGFVACLRRLVPPGLWGS 502
|||||
Db 481 RHNERFLRNTKFFISLGHAKLSLQELTWKMSVRDCAMLRSPGVCVPAASHRLREEI 540
|||||
QY 503 RHNERFLRNTKFFISLGHAKLSLQELTWKMSVRDCAMLRSPGVCVPAASHRLREEI 562
|||||
Db 541 LAKFLHMLSVYVVELLRSGFFVYTTTFQKNRLLFFYKRSVWSKLSQIGIRQHLKRVQLRE 600
|||||
QY 563 LAKFLHMLSVYVVELLRSGFFVYTTTFQKNRLLFFYKRSVWSKLSQIGIRQHLKRVQLRE 622
|||||
Db 601 LSAEVRQHRREARPAALLTSKRLRIPKPDGLRPIVNMMDYVVGARTFRREKRAERLTSRKA 660
|||||
QY 623 LSAEVRQHRREARPAALLTSKRLRIPKPDGLRPIVNMMDYVVGARTFRREKRAERLTSRKA 682
|||||
Db 661 LFSVLNVERARRPGLLGASVIGLDDIHRARWTEFLVRADQDPPPELYFVKVDVTGAYDTI 720
|||||
QY 683 LFSVLNVERARRPGLLGASVIGLDDIHRARWTEFLVRADQDPPPELYFVKVDVTGAYDTI 742
|||||
Db 721 PODRLTEVIASIIKPQNTYCVRRYAVVQKAAHGHVKAFKSHVSTLTLDLPYMRQFVAHL 780
|||||
QY 743 PODRLTEVIASIIKPQNTYCVRRYAVVQKAAHGHVKAFKSHVSTLTLDLPYMRQFVAHL 802
|||||
Db 781 QETSPLEADAVVIEQSSSLNEASSGLFDVFLRFMCHHAVIRGKSVYVQCQIGPOGSILSTL 840

QY 803 QETSPRLDAWVIEQSSSLNEASSGLEDVFLFENCHAVIRKSYVQCOGIFQGSILSTL 862
DB 841 LCSLCYGDMEKLFAGIRRDGLLLRLVDVDFLLVPHLTTHAKTFLRLVGRVPEYGCVVNL 900
QY 863 LCSLCYGDMEKLFAGIRRDGLLLRLVDVDFLLVPHLTTHAKTFLRLVGRVPEYGCVVNL 922
DB 901 RKTUVNFPVEDEALGCTAFVQMPAHGLFPWCGLLDTRILEVQSDYSYVARTSIRASLTF 960
QY 923 RKTUVNFPVEDEALGCTAFVQMPAHGLFPWCGLLDTRILEVQSDYSYVARTSIRASLTF 982
DB 961 NRGFKAGRNRRKLFVLRKCHSLFELDLQVNSLQVCTNIYKILLQAYRHACVQLP 1020
QY 983 NRGFKAGRNRRKLFVLRKCHSLFELDLQVNSLQVCTNIYKILLQAYRHACVQLP 1042
DB 1021 FHOQVWKNPTFFLRVISTASLCYSILKAKNAGMSLGAKGAGPLPSEAVQWLCHQAFLL 1080
QY 1043 FHOQVWKNPTFFLRVISTASLCYSILKAKNAGMSLGAKGAGPLPSEAVQWLCHQAFLL 1102
DB 1081 KLTRHRTVYVPLGLSLRTAQTLQSRKLPGLTTLTALEAANPALPSDFKTILD 1132
QY 1103 KLTRHRTVYVPLGLSLRTAQTLQSRKLPGLTTLTALEAANPALPSDFKTILD 1154

RESULT 15
ID US-08-912-951-2 STANDARD; PRT; 1132 AA.

AC xxxxxx

Sequence 2, Application US/08912951

Sequence 2, Application US/08912951

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin

APPLICANT: Andrews, William H.

TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND

TITLE OF INVENTION: THERAPEUTIC METHODS

NUMBER OF SEQUENCES: 335

CORRESPONDENCE ADDRESS:

ADDRESS: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/912,951

FILING DATE: 14-AUG-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

CLASSIFICATION: 435

CC

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-0026000US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1132 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1132 AA; 126995 MW; 6588209 CN;

Query Match 98.2%; Score 8465; DB 13; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MPAPRCRAVRSLLRSHYREVLPATFVRRLLGPGQWRLLVQRGDPAAFRALVAQCILVCVPW 60
QY 23 MPAPRCRAVRSLLRSHYREVLPATFVRRLLGPGQWRLLVQRGDPAAFRALVAQCILVCVPW 82

DB 61 DARPPPAAPSFQVSCLELVARVQLRCERGAKNVLAFFGALLDARGGPPPEATTTSVR 120
QY 83 DARPPPAAPSFQVSCLELVARVQLRCERGAKNVLAFFGALLDARGGPPPEATTTSVR 142

DB 121 SYLPNTVTDALRGSGAWGLLRVGGDVLVLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
QY 143 SYLPNTVTDALRGSGAWGLLRVGGDVLVLLARCALFVLVAPSCAYQVCGPPLYQLGA 202

DB 181 ATQARPPPHASGPRRRRLGCRANWHSVREAGVPLGLPAPGARRRGGSASRSPLPKRPRR 240
QY 203 ATQARPPPHASGPRRRRLGCRANWHSVREAGVPLGLPAPGARRRGGSASRSPLPKRPRR 262

DB 241 GAAPPERTPVQGSWAHPGTRGPDSDRGFCVVSAPARPAEATSLGALSGTRHSHPSVG 300
QY 263 GAAPPERTPVQGSWAHPGTRGPDSDRGFCVVSAPARPAEATSLGALSGTRHSHPSVG 322

DB 301 RQHAGPPSTSRPPRPWDTCPPPVYAEKHFYSSGDKQLRPSFLLSLRSLTGARRL 360
QY 323 RQHAGPPSTSRPPRPWDTCPPPVYAEKHFYSSGDKQLRPSFLLSLRSLTGARRL 382

DB 361 VETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPLRAAVT 420
QY 383 VETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPLRAAVT 442

DB 421 PAAGVAREKPGQSVAAPEEEDTPRRLVQLRQHSQWVYGFVRACLRRLVPPGLMGS 480
QY 443 PAAGVAREKPGQSVAAPEEEDTPRRLVQLRQHSQWVYGFVRACLRRLVPPGLMGS 502

DB 481 RNERREFLNTKKFISLGHAKLSLQELTWKMSVRDCAWLRSRSPGVGCVPAAEHLRREI 540
QY 503 RNERREFLNTKKFISLGHAKLSLQELTWKMSVRDCAWLRSRSPGVGCVPAAEHLRREI 562

DB 541 LAKFLHMLMSVYVVELLSRFFVYTTTFOKNRLFFYRKSWSKLSIGIRQHLKRVQLRE 600
QY 563 LAKFLHMLMSVYVVELLSRFFVYTTTFOKNRLFFYRKSWSKLSIGIRQHLKRVQLRE 622

DB 601 LSEAEVQRHREARPAALLTSRLRFIPKPDGLRPIVNMVYVVGARTFRREKRAERLTSRYKA 660
QY 623 LSEAEVQRHREARPAALLTSRLRFIPKPDGLRPIVNMVYVVGARTFRREKRAERLTSRYKA 682

DB 661 LFSVLNFERARRPGLLGASVLGLDDIHRAWRTFVLVRAQDPPPELYFKVDVTGAYDTI 720
|||||

QY 683 LFSVLNVERARRRGLLGASVLGDDIHRWRTFVLVRQAQDPPPELYFVKVDVTGAYDI 742
Db 721 PODRLTEVIASIIKQNTYICVRRYAVVQKAAGHVRKAFKSHYSTLTDLQPYMRQFVAHL 780
QY 743 PODRLTEVIASIIKQNTYICVRRYAVVQKAAGHVRKAFKSHYSTLTDLQPYMRQFVAHL 802
Db 781 QETSPRLADAVVIEQSSSLNEASSGLFDFVLFKCHHVRIRGKSYVQCQIGPOGSIILSTL 840
QY 803 QETSPRLADAVVIEQSSSLNEASSGLFDFVLFKCHHVRIRGKSYVQCQIGPOGSIILSTL 862
Db 841 LCSLCYGDMEKLFAGIRRRDGLLRVDDFLLVTPHLLTHAKTFLRLTVRGVPEYGCVVNL 900
QY 863 LCSLCYGDMEKLFAGIRRRDGLLRVDDFLLVTPHLLTHAKTFLRLTVRGVPEYGCVVNL 922
Db 901 RKTWNFPVEDEALGTAFTQMPAHGLFPMCGLLDTRTLEVOQSDYSSVARTSIRASLTF 960
QY 923 RKTWNFPVEDEALGTAFTQMPAHGLFPMCGLLDTRTLEVOQSDYSSVARTSIRASLTF 982
Db 961 NRGFKAGRNRRKLFGLVRLKCHSLFELDLQVNSLQTVCTNIYKILLQAYRFHACVQLP 1020
QY 983 NRGFKAGRNRRKLFGLVRLKCHSLFELDLQVNSLQTVCTNIYKILLQAYRFHACVQLP 1042
Db 1021 FHOQWKNPTFFELRVISDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVQWLCHQAFLL 1080
QY 1043 FHOQWKNPTFFELRVISDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVQWLCHQAFLL 1102
Db 1081 KLTRHRVTVYVPLGSLRTAQTLQSRKLPGLTTLTALEAAANPALPSDFKTLTD 1132
QY 1103 KLTRHRVTVYVPLGSLRTAQTLQSRKLPGLTTLTALEAAANPALPSDFKTLTD 1154

RESULT 16
ID US-08-911-312-2 STANDARD: PRT: 1132 AA.
XX
AC
XX
XX
DT
XX

Sequence 2, Application US/08911312

Sequence 2, Application US/08911312

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William

TITLE OF INVENTION: Telomerase Reverse Transcriptase

NUMBER OF SEQUENCES: 170

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/911,312

FILING DATE: 14-AUG-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

CG

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Einhorn, Gregory P.
CC REGISTRATION NUMBER: 38,440
CC REFERENCE/DOCKET NUMBER: 015389-0025000US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1132 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE: 1132 AA; 126995 MW; 6588209 CN;

Query Match 98.2%; Score 8465; DB 13; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MPRAPRCRAVRSLLRSHYREVLPATFVRRLLGPQWRVLVQRGDPAAFRALVAQCLVCPW 60
QY 23 MPRAPRCRAVRSLLRSHYREVLPATFVRRLLGPQWRVLVQRGDPAAFRALVAQCLVCPW 82
Db 61 DARPPPAAPSFQVSCLELVARVQLRCERGAKNVLAFFGALLDARGGPPPEATTTSVR 120
QY 83 DARPPPAAPSFQVSCLELVARVQLRCERGAKNVLAFFGALLDARGGPPPEATTTSVR 142
Db 121 SYLPNTVTDALRGSGAWGLLRVGDVLLHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
QY 143 SYLPNTVTDALRGSGAWGLLRVGDVLLHLLARCALFVLVAPSCAYQVCGPPLYQLGA 202
Db 181 ATQARPPPHASCPRRRLGGERAWNHSVREAGVPLGLPAGARRRGGSASRSLPLKRRPR 240
QY 203 ATQARPPPHASCPRRRLGGERAWNHSVREAGVPLGLPAGARRRGGSASRSLPLKRRPR 262
Db 241 GAAPEPRTPVQGSWAHPGRTGRGSDRGFCVVSAPAEAEATSLGALSGTRHSHPSVG 300
QY 263 GAAPEPRTPVQGSWAHPGRTGRGSDRGFCVVSAPAEAEATSLGALSGTRHSHPSVG 322
Db 301 ROHAGPPSTSRPPRPWDTPCPVYAETKHFYSSGDEQLRPSFLLSLRSLTGARRL 360
QY 323 ROHAGPPSTSRPPRPWDTPCPVYAETKHFYSSGDEQLRPSFLLSLRSLTGARRL 382
Db 361 VETIFLGSRRPMGTPRRLPLRPLQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPRAAVT 420
QY 383 VETIFLGSRRPMGTPRRLPLRPLQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPRAAVT 442
Db 421 PAAGVCAREKPOGSVAAPDEEDTPRRLVQLLRQHSHPQVYGFVRACLRRLRVPPLMGWS 480
QY 443 PAAGVCAREKPOGSVAAPDEEDTPRRLVQLLRQHSHPQVYGFVRACLRRLRVPPLMGWS 502
Db 481 RHNRERFLNRTKKFISLGKHAQLSLQELTWMKSVRDCAWLRSPGVGCVPAAEHLRREI 540
QY 503 RHNRERFLNRTKKFISLGKHAQLSLQELTWMKSVRDCAWLRSPGVGCVPAAEHLRREI 562
Db 541 LAKFLHLMMSVYVVELLSFFVYTTTQKNEFLFYRKSWSKLSQISGIRQHLKRVOLRE 600
QY 563 LAKFLHLMMSVYVVELLSFFVYTTTQKNEFLFYRKSWSKLSQISGIRQHLKRVOLRE 622

Db 191 ATQARPPPHASGRRRLGGERAWNHVSREAGVPLGLPAGARRRGGASRSLPLPKRPRR 240
Qy 203 ATQARPPPHASGRRRLGGERAWNHVSREAGVPLGLPAGARRRGGASRSLPLPKRPRR 262
Db 241 GAAPERTPVGQSWAHPGTRGSDRGFCVSPARPAEATSELEGALSCTRSHPSVG 300
Qy 263 GAAPERTPVGQSWAHPGTRGSDRGFCVSPARPAEATSELEGALSCTRSHPSVG 322
Db 301 ROHAGPSTSRPRPMDTPCPVYAEKHELYSSGKEQLRPSFLSSLRPSLTGARRL 360
Qy 323 ROHAGPSTSRPRPMDTPCPVYAEKHELYSSGKEQLRPSFLSSLRPSLTGARRL 382
Db 361 VETIFLGSRPWMPGTRPRLPLRPQRYWQMRPLFLELGNHQAQCYGVLLKTHCPLRAAVT 420
Qy 383 VETIFLGSRPWMPGTRPRLPLRPQRYWQMRPLFLELGNHQAQCYGVLLKTHCPLRAAVT 442
Db 421 PAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRHSSPWQYGVFRACLRRLVPPGLWGS 480
Qy 443 PAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRHSSPWQYGVFRACLRRLVPPGLWGS 502
Db 481 RHNERFLRNTKKTISLGKHAHKLQELTWKMSVRDCAWLRSPGVCVPAAEHRLREEI 540
Qy 503 RHNERFLRNTKKTISLGKHAHKLQELTWKMSVRDCAWLRSPGVCVPAAEHRLREEI 562
Db 541 LAKELHMLMSVYVELLSRFFYVTTETFPQKNRLFYRKSVMSKLSQSIGRHLKRVQVLE 600
Qy 563 LAKELHMLMSVYVELLSRFFYVTTETFPQKNRLFYRKSVMSKLSQSIGRHLKRVQVLE 622
Db 601 LSEAEVROHREARPAALLTSRLRFPKPDGLRPIVNMVYVVGARTFREKRAERLTSRVKA 660
Qy 623 LSEAEVROHREARPAALLTSRLRFPKPDGLRPIVNMVYVVGARTFREKRAERLTSRVKA 682
Db 661 LFSVLNTERARRPGLLGASVLGLDDIHRARTFVLVRAQDPPPELVKVDVTGAYDTI 720
Qy 683 LFSVLNTERARRPGLLGASVLGLDDIHRARTFVLVRAQDPPPELVKVDVTGAYDTI 742
Db 721 PODRLTEVIASIIKPQNTYCVRRYAVVOKAAGHVRKAFKSHVSTLTDLQPYMRQFVAHL 780
Qy 743 PODRLTEVIASIIKPQNTYCVRRYAVVOKAAGHVRKAFKSHVSTLTDLQPYMRQFVAHL 802
Db 781 QETSPRLDAVVEQSSLINEASSGLFVFLRPMCHHVRIRGKSVYVOCQIPQGSILSTL 840
Qy 803 QETSPRLDAVVEQSSLINEASSGLFVFLRPMCHHVRIRGKSVYVOCQIPQGSILSTL 862
Db 841 LCSLCYGDMEKLFAGIRRDGILLRLVDDFLVTPHLLTHAKTFLRTLVRGYPEYGVVNL 900
Qy 863 LCSLCYGDMEKLFAGIRRDGILLRLVDDFLVTPHLLTHAKTFLRTLVRGYPEYGVVNL 922
Db 901 RKTVVNPFVEALGGTAFVQMPAHGLFPWCGLLDDTLRTLEVQSDYSYARTSTRASLTF 960
Qy 923 RKTVVNPFVEALGGTAFVQMPAHGLFPWCGLLDDTLRTLEVQSDYSYARTSTRASLTF 982
Db 961 NRGFKAGNMRKLFGLVRLKCHSLFDLVQNSLQTVCTNLYKILLQAYRFHACVQLP 1020
Qy 983 NRGFKAGNMRKLFGLVRLKCHSLFDLVQNSLQTVCTNLYKILLQAYRFHACVQLP 1042
Db 1021 FHQWKNPTFFLRVYISDTASLCYSILKAKNAGMSLGAAGPLPSEAVOWLCHQAFLL 1080
Qy 1043 FHQWKNPTFFLRVYISDTASLCYSILKAKNAGMSLGAAGPLPSEAVOWLCHQAFLL 1102
Db 1081 KLTHRVYVPLLSLRTAQQLSRKPLGPTLTALAAANPALPSDFKTIILD 1132
Qy 1103 KLTHRVYVPLLSLRTAQQLSRKPLGPTLTALAAANPALPSDFKTIILD 1154

RESULT 18
ID US-09-052-919-2 STANDARD: PRT; 1132 AA.
XX
AC xxxxxx
XX
DT
XX.

DE Sequence 2, Application US/09052919
XX Sequence 2, Application US/09052919
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Antisense Compositions for Detecting and
CC NUMBER OF SEQUENCES: 72
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/052,919
CC FILING DATE: 31-MAR-1998
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/974,549
CC FILING DATE: 19-NOV-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/974,584
CC FILING DATE: 19-NOV-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Parent, Annette S.
CC REGISTRATION NUMBER: 42,058
CC REFERENCE/DOCKET NUMBER: 015389-0036000US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300

CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1132 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1132 AA; 126995 MW; 6588209 CN;

Query Match 98.28; Score 8465; DB 14; Length 1132;
Best Local Similarity 100.08; Pred. No. 0.00e+00;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MPAPRCRAVRSLLRSHYREVLPVRLGQWRLVORGDPAAFRALVAQCLVCPW 60
QY 23 MPAPRCRAVRSLLRSHYREVLPVRLGQWRLVORGDPAAFRALVAQCLVCPW 82
Db 61 DARPPPAAPSFQVSCLELVARVLQRCERGAKNVLAFAFDLGARGGPPPEAFTTSVR 120
QY 83 DARPPPAAPSFQVSCLELVARVLQRCERGAKNVLAFAFDLGARGGPPPEAFTTSVR 142
Db 121 SYLPNTVTDALRGSGAGLLRLRRVDDVLLHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
QY 143 SYLPNTVTDALRGSGAGLLRLRRVDDVLLHLLARCALFVLVAPSCAYQVCGPPLYQLGA 202
Db 181 ATQARPPHAGSRRRLGRCERANNHVSREAGVPLGIPAPGARRRGGASRSLLPLKPRRR 240
QY 203 ATQARPPHAGSRRRLGRCERANNHVSREAGVPLGIPAPGARRRGGASRSLLPLKPRRR 262
Db 241 GAAPEPERTVPGOGSWAHPGRTGSPDRGFCVVSPPARPAEATSEALSGTRHSHPSVG 300
QY 263 GAAPEPERTVPGOGSWAHPGRTGSPDRGFCVVSPPARPAEATSEALSGTRHSHPSVG 322
Db 301 RQHAGPPSTSRPPRPWDTPCPVPVYAEKHFYSSGDKQLRPSFLSLRPSLTGARRL 360
QY 323 RQHAGPPSTSRPPRPWDTPCPVPVYAEKHFYSSGDKQLRPSFLSLRPSLTGARRL 382
Db 361 VETIFLGSRRPMTGTPRRLPRLPQRYQWMPLELLELGNHAQCPYGVLLTHCPRAAVT 420
QY 383 VETIFLGSRRPMTGTPRRLPRLPQRYQWMPLELLELGNHAQCPYGVLLTHCPRAAVT 442
Db 421 PAAGVCAREKPOGSVAAPAEEDTPRLVOLLROHSSPMQVYGFVRACLRRLLVPPGLWGS 480
QY 443 PAAGVCAREKPOGSVAAPAEEDTPRLVOLLROHSSPMQVYGFVRACLRRLLVPPGLWGS 502
Db 481 RHNERELRNTKRFISLGHAKLSLQELTWKMSVRDCAWLRRSPGVGCVPAAEHRLREEI 540
QY 503 RHNERELRNTKRFISLGHAKLSLQELTWKMSVRDCAWLRRSPGVGCVPAAEHRLREEI 562
Db 541 LAKFLHLMSSVYVVELLSRFYVTTETFOKNRLFYRKSWKLSQSGIRQHLKRVOLRE 600
QY 563 LAKFLHLMSSVYVVELLSRFYVTTETFOKNRLFYRKSWKLSQSGIRQHLKRVOLRE 622
Db 601 LSAEVRQHREARPAALLTSRLRFTPKDGLRPIVNMDDYVVGARTFREKRAERLTSRVKA 660
QY 623 LSAEVRQHREARPAALLTSRLRFTPKDGLRPIVNMDDYVVGARTFREKRAERLTSRVKA 682
Db 661 LFSVLNTERARRPGLLGASVLGDDIHRANWTFVLRVRAQDPPPELVFVKVDVTGAYDTI 720
QY 683 LFSVLNTERARRPGLLGASVLGDDIHRANWTFVLRVRAQDPPPELVFVKVDVTGAYDTI 742
Db 721 PQRLTEVIASIKPQNTYCVRRYAVVQKAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 780
QY 743 PQRLTEVIASIKPQNTYCVRRYAVVQKAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 802
Db 781 QETSPLRDADVIEQSSSLNEASSGLFDVLFREMHCHAVRIGKSYVOCQIPOGSILSTL 840
QY 803 QETSPLRDADVIEQSSSLNEASSGLFDVLFREMHCHAVRIGKSYVOCQIPOGSILSTL 862
Db 841 LCSICYGDMENKLFAGIRRDGLLRLVDDFLVTPHLLTHAKTFLRTLVRGVPYEGCVVNL 900
QY 863 LCSICYGDMENKLFAGIRRDGLLRLVDDFLVTPHLLTHAKTFLRTLVRGVPYEGCVVNL 922
Db 901 RKTVVNFPVEDEALGGTAFVQMPAHGLFPWCGLLDTRTLEVQSDYSYARTSIRASLTF 960

QY 923 RKTVVNFPVEDEALGGTAFVQMPAHGLFPWCGLLDTRTLEVQSDYSYARTSIRASLTF 982
Db 961 NRGFKAGRNRRRLFGVLRLLKCHSLFILDQVNSIQVCTNYIKLLIQAYRFHACVQLQP 1020
QY 983 NRGFKAGRNRRRLFGVLRLLKCHSLFILDQVNSIQVCTNYIKLLIQAYRFHACVQLQP 1042
Db 1021 FHOQVKNPTFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVQWLCHQAFLL 1080
QY 1043 FHOQVKNPTFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVQWLCHQAFLL 1102
Db 1081 KLFRHRVTVYVPLGLSLRTAQOTLSRKLPGTTLTAAEAANPALPSDFKTILD 1132
QY 1103 KLFRHRVTVYVPLGLSLRTAQOTLSRKLPGTTLTAAEAANPALPSDFKTILD 1154
RESULT 19
ID US-08-854-050-225 STANDARD; PRT; 1132 AA.
XX
AC xxxxxx
XX
DT
XX
DE
XX
CC Sequence 225, Application US/08854050
CC Sequence 225, Application US/08854050
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Novel Telomerase
CC NUMBER OF SEQUENCES: 225
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002930US
CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 225:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1132 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 1132 AA: 126995 MW: 6588209 CN;

Query Match 98.2%; Score 8465; DB 12; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MPRAPRCRAVRLSHRYREVLPVPLATFVRRLGPOGWRVLVORGDPAAFRALVAOCLVCPW 60
QY 23 MPRAPRCRAVRLSHRYREVLPVPLATFVRRLGPOGWRVLVORGDPAAFRALVAOCLVCPW 82

Db 61 DAREPPAAPSPFROVSCUKELVARVLQRLCERGAKNVLAFFALLDGGARGGPPEAFTTSVR 120
QY 83 DAREPPAAPSPFROVSCUKELVARVLQRLCERGAKNVLAFFALLDGGARGGPPEAFTTSVR 142

Db 121 SYLPTVTDALRGSGANGLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLYOLGA 180
QY 143 SYLPTVTDALRGSGANGLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLYOLGA 202

Db 181 ATQARPPHAGSPRRRLGCRANMHSVREAGVPLGAPGARRGGASRLPLPKRPRR 240
QY 203 ATQARPPHAGSPRRRLGCRANMHSVREAGVPLGAPGARRGGASRLPLPKRPRR 262

Db 241 GAPEPERTVPGQSWAHPTGRTGSPDRGFCVVSPPARPAEATSLGALSSTRHSPVSG 300
QY 263 GAPEPERTVPGQSWAHPTGRTGSPDRGFCVVSPPARPAEATSLGALSSTRHSPVSG 322

Db 301 ROHAGPPTSRRPPWDTPCPVPYATTKHFLYSSGKEQLRPSFLSSLRPSLTGARRL 360
QY 323 ROHAGPPTSRRPPWDTPCPVPYATTKHFLYSSGKEQLRPSFLSSLRPSLTGARRL 382

Db 361 VETIFLSRPWMPGTGTRRLPLRQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPRAAVT 420
QY 383 VETIFLSRPWMPGTGTRRLPLRQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPRAAVT 442

Db 421 PAAGVCAREKPOGSVAPEEEDTPRLVOLLRHSSPMQVYGFVRACTRLRPLPGLWGS 480
QY 443 PAAGVCAREKPOGSVAPEEEDTPRLVOLLRHSSPMQVYGFVRACTRLRPLPGLWGS 502

Db 481 RHNERRELRNTKRFISLGGKAKLSLQELTWKMSVRDCAWLRSPGVCVPAAEHRLREEI 540
QY 503 RHNERRELRNTKRFISLGGKAKLSLQELTWKMSVRDCAWLRSPGVCVPAAEHRLREEI 562

Db 541 LAKFLHLMVSVYVVELLRSFFYVTTFTQKNRLFYFKRKSVMKLSQSTIGIROHLKRVLRE 600
QY 563 LAKFLHLMVSVYVVELLRSFFYVTTFTQKNRLFYFKRKSVMKLSQSTIGIROHLKRVLRE 622

Db 601 LSEAEVQHEAREAPALTSRLRFTPKDGLRPVNMVYVVGATFREKRAERLTSRVKA 660
QY 623 LSEAEVQHEAREAPALTSRLRFTPKDGLRPVNMVYVVGATFREKRAERLTSRVKA 682

Db 661 LFSVLNTERARRPGLLGASVLGDDIHRAWRTFVLVRAODPPPELVFVKVDVTGAYDTI 720
QY 683 LFSVLNTERARRPGLLGASVLGDDIHRAWRTFVLVRAODPPPELVFVKVDVTGAYDTI 742

Db 721 PQDLRTVIAIITKPNQYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 780
QY 743 PQDLRTVIAIITKPNQYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 802

Db 781 QETSPRLDAVVEQSSSLNEASSGLDFVLRFMCHHAVIRGKSYVQCQIPQGSILSTL 840
QY 803 QETSPRLDAVVEQSSSLNEASSGLDFVLRFMCHHAVIRGKSYVQCQIPQGSILSTL 862

Db 841 LCSLCYGDMEKNKLFAGIRRRDGLLRLVDDFLVTPHLLTHAKTFLRTLVRGVPYGCVVNL 900
QY 863 LCSLCYGDMEKNKLFAGIRRRDGLLRLVDDFLVTPHLLTHAKTFLRTLVRGVPYGCVVNL 922

Db 901 RKTVVNFPVEDEALGCTAFVQMPAHGLFPWCGLLLDTRILEVQSDYSSYARTSIRASLTF 960
QY 923 RKTVVNFPVEDEALGCTAFVQMPAHGLFPWCGLLLDTRILEVQSDYSSYARTSIRASLTF 982

Db 961 NRGFKAGRNMRRLFGVLRKCHSLFELDQVNSLOTVCTNIYKILLQAYRFHACVQLP 1020
QY 983 NRGFKAGRNMRRLFGVLRKCHSLFELDQVNSLOTVCTNIYKILLQAYRFHACVQLP 1042

Db 1021 FHQOVKNPTFFFLRVISDTASLCYSILKAKNAGMSLGAKGAGPLPSEAVQWMLCHQAFLL 1080
QY 1043 FHQOVKNPTFFFLRVISDTASLCYSILKAKNAGMSLGAKGAGPLPSEAVQWMLCHQAFLL 1102

Db 1081 KLTRHRVTYVPLGSLRTAQTLRSKLPGLTTLTALEAAANPALPSDFKTILD 1132
QY 1103 KLTRHRVTYVPLGSLRTAQTLRSKLPGLTTLTALEAAANPALPSDFKTILD 1154

RESULT 20
ID US-08-912-951-323 STANDARD; PRT; 1154 AA.
XX
AC xxxxxx
DT
XX
DT
XX
DE
XX
Sequence 323, Application US/08912951
Sequence 323, Application US/08912951
GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC TITLE OF INVENTION: THERAPEUTIC METHODS
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent in Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643

CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002600US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 323:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1154 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 1154 AA; 129603 MW; 6816882 CN;

Query Match 98.2%; Score 8465; DB 13; Length 1154;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRLVQRGDPAAFRALVAQCLVCVW 60
QY 23 MPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRLVQRGDPAAFRALVAQCLVCVW 82

Db 61 DARPPPAAPSFQVSCLELVARVQLRCERGAKNVLAFFGALLDGAARGPPPEAFTTSVR 120
QY 83 DARPPPAAPSFQVSCLELVARVQLRCERGAKNVLAFFGALLDGAARGPPPEAFTTSVR 142

Db 121 SYLPNTVTDLRGSGAWGLLRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
QY 143 SYLPNTVTDLRGSGAWGLLRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 202

Db 181 ATQARPPPHASGPRRLGGERAWNVSRVAGVPLGLPAPGARRRGGSASRSLPLPKRRR 240
QY 203 ATQARPPPHASGPRRLGGERAWNVSRVAGVPLGLPAPGARRRGGSASRSLPLPKRRR 262

Db 241 GAAPERTPVGGGWAHPGRTGSDRGFCVVSPARPAEATSLBGALSGTRHSHPSVG 300
QY 263 GAAPERTPVGGGWAHPGRTGSDRGFCVVSPARPAEATSLBGALSGTRHSHPSVG 322

Db 301 ROHAGPSTSRPPRWDTPCPVYAEKHFYSSGDKQLRPSFLLSLRSLTGARL 360
QY 323 ROHAGPSTSRPPRWDTPCPVYAEKHFYSSGDKQLRPSFLLSLRSLTGARL 382

Db 361 VETIFGSRPMPGTPRLPLPQRYWQMRPLFLELLGHNACPYGVLLKTHCPLRAAVT 420
QY 383 VETIFGSRPMPGTPRLPLPQRYWQMRPLFLELLGHNACPYGVLLKTHCPLRAAVT 442

Db 421 PAAGVCAREKPGQSVAAPEEDTPRRLVQLLRQHSPPQVYGFVRACLRLRVPPLGWS 480
QY 443 PAAGVCAREKPGQSVAAPEEDTPRRLVQLLRQHSPPQVYGFVRACLRLRVPPLGWS 502

Db 481 RHNERFLNWKFFISLGHAKLSLQELTWKMSVRDCAWLRRSPGVCVPAAEHRUREEI 540
QY 503 RHNERFLNWKFFISLGHAKLSLQELTWKMSVRDCAWLRRSPGVCVPAAEHRUREEI 562

Db 541 LAKFLHLMMSVYVVELLSFFVYTTTFQKNRLLFFYKRSWWSKLSQIGIRHLLKRVQLRE 600
QY 563 LAKFLHLMMSVYVVELLSFFVYTTTFQKNRLLFFYKRSWWSKLSQIGIRHLLKRVQLRE 622

Db 601 LSEAEVQHQREARPAALLTSRLRFIPKPDGLRPIVNDVYVGARTFRREKRAELTSRVKA 660
QY 623 LSEAEVQHQREARPAALLTSRLRFIPKPDGLRPIVNDVYVGARTFRREKRAELTSRVKA 682

Db 661 LFSVLNFERARRPGLLGASVLGLDDTHRAWRFTVLVRAQDPPPELYFVKVDVTGAYDTI 720
QY 683 LFSVLNFERARRPGLLGASVLGLDDTHRAWRFTVLVRAQDPPPELYFVKVDVTGAYDTI 742

Db 721 PDRLTEVIASIIKPNQTCVRYAVVQKAAHGKVKAFKSHVSTLTDLQPYMRQFVAHL 780
QY 743 PDRLTEVIASIIKPNQTCVRYAVVQKAAHGKVKAFKSHVSTLTDLQPYMRQFVAHL 802

Db 781 QETSPLRDAVVIQSSSLNEASSGLFDVFLRFMCHHAVIRKSVYQCOGIPQGSILSTL 840
QY 803 QETSPLRDAVVIQSSSLNEASSGLFDVFLRFMCHHAVIRKSVYQCOGIPQGSILSTL 862

Db 841 LCSLCYGDMMENKLFAGIRRDGLLLRLVDDFLVTLPHLTHAKTFLRTLVRGVEYGCVVNL 900
QY 863 LCSLCYGDMMENKLFAGIRRDGLLLRLVDDFLVTLPHLTHAKTFLRTLVRGVEYGCVVNL 922

Db 901 RKTVVNFPVEDEALGTAFFVQMPAHGLFPWCGLLLDTRTLEVQSDYSYVARTSIRASLTF 960
QY 923 RKTVVNFPVEDEALGTAFFVQMPAHGLFPWCGLLLDTRTLEVQSDYSYVARTSIRASLTF 982

Db 961 NRGFRAGRNMRKLFGLVLRKCHSLFLDLQVNSLOTVCNIIKILLQAYRHACVLQLP 1020
QY 983 NRGFRAGRNMRKLFGLVLRKCHSLFLDLQVNSLOTVCNIIKILLQAYRHACVLQLP 1042

Db 1021 FHQOVWKNPTFFELRVISDTASLCYSTLKAKNAGMSLGAKGAGPLPSEAVQWLCHQAFLL 1080
QY 1043 FHQOVWKNPTFFELRVISDTASLCYSTLKAKNAGMSLGAKGAGPLPSEAVQWLCHQAFLL 1102

Db 1081 KLTRHRVTYVPLLGSLRTAQTLRSRKLPGTTLTALAAANPALPSDFKTILD 1132
QY 1103 KLTRHRVTYVPLLGSLRTAQTLRSRKLPGTTLTALAAANPALPSDFKTILD 1154

RESULT 21
ID US-08-974-549-611 STANDARD; PRT: 1154 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 611, Application US/08974549
XX
CC Sequence 611, Application US/08974549
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Human Telomerase Catalytic Subunit
CC NUMBER OF SEQUENCES: 726
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/974,549
CC FILING DATE: 19-NOV-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843

CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph Ted
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002610US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 611:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1154 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: Protein
CC LOCATION: 1..1154
CC OTHER INFORMATION: /note= "fusion protein composed of hTERT
CC OTHER INFORMATION: protein sequence, vector sequences, the
CC OTHER INFORMATION: Myc epitope and His6 tag"
CC SEQUENCE 1154 AA; 129603 MW; 6816882 CN;
SQ
Query Match 98.28; Score 8465; DB 18; Length 1154;
Best Local Similarity 100.08; Pred. No. 0.00e+00;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 MPAPRCRAVRSLLRSHYREVLPATFVRLGPGQWRLVQSGDPAAFRALVAQCLVCPW 60
QY 23 MPAPRCRAVRSLLRSHYREVLPATFVRLGPGQWRLVQSGDPAAFRALVAQCLVCPW 82
Db 61 DARPPPAAPSFQVSCLELVARVQLRCERGAKNVLAFGFALLDGAAGGPPPEAFTTSVR 120
QY 83 DARPPPAAPSFQVSCLELVARVQLRCERGAKNVLAFGFALLDGAAGGPPPEAFTTSVR 142
Db 121 SYLPNTVTDALRGSGAWGLLRVGGDVLVLLARCALFVLVAPSCAYQVCGPPLYLGA 180
QY 143 SYLPNTVTDALRGSGAWGLLRVGGDVLVLLARCALFVLVAPSCAYQVCGPPLYLGA 202
Db 181 ATQARPPPHASGRRRLGGERANVHSYREAGVPLGLPAPGARRRGSGASRLPLKRRPR 240
QY 203 ATQARPPPHASGRRRLGGERANVHSYREAGVPLGLPAPGARRRGSGASRLPLKRRPR 262
Db 241 GAAPERTPVGGSWAHGPTRGSDRGFCVVSYPARPAEATSLGALSGTRHSPSVG 300
QY 263 GAAPERTPVGGSWAHGPTRGSDRGFCVVSYPARPAEATSLGALSGTRHSPSVG 322
Db 301 RQHHAGPPTSRRPMDTCCPVYATKHFLLYSSGDKQLRFSLLSLRPLSGARLL 360
QY 323 RQHHAGPPTSRRPMDTCCPVYATKHFLLYSSGDKQLRFSLLSLRPLSGARLL 382
Db 361 VETIFLGSRPMPGTPRRRLPRLPQRYWQMPLELLGNHAQCPYGVLLKTHCPRAAVT 420
QY 383 VETIFLGSRPMPGTPRRRLPRLPQRYWQMPLELLGNHAQCPYGVLLKTHCPRAAVT 442

Db 421 PAAGVCAREKPOGSGVAAPDEEDTDPRLVQLLRQHSPPQVYGFVACLRRLVPPGLWGS 480
QY 443 PAAGVCAREKPOGSGVAAPDEEDTDPRLVQLLRQHSPPQVYGFVACLRRLVPPGLWGS 502
Db 481 RNREFLRNTKFFISLGHAKLSLQELTWKMSVRDCAWLRSPGVGCVPAAEHRLREI 540
QY 503 RNREFLRNTKFFISLGHAKLSLQELTWKMSVRDCAWLRSPGVGCVPAAEHRLREI 562
Db 541 LAKFLHMLMSVVVVELLSFFVYVTTTFQKNRLLFFYKSWKSLQSIGIRQHLKRVQRE 600
QY 563 LAKFLHMLMSVVVVELLSFFVYVTTTFQKNRLLFFYKSWKSLQSIGIRQHLKRVQRE 622
Db 601 LSEAEVRQHREARPAALLTSRLRFIPKPDGLRPVNNMDYVVGARTFRREKRAERLTSRKA 660
QY 623 LSEAEVRQHREARPAALLTSRLRFIPKPDGLRPVNNMDYVVGARTFRREKRAERLTSRKA 682
Db 661 LFSVLNRYERARRPGLLGASVLGDDIHRARWTFVLRVRAODPPPELYFYKVDVTGAYDI 720
QY 683 LFSVLNRYERARRPGLLGASVLGDDIHRARWTFVLRVRAODPPPELYFYKVDVTGAYDI 742
Db 721 PQDRLTEVIASIIKQNTYCVRRYAVVQKAAHGHVKAFAKSHVSTLTDLPYMRQFVAHL 780
QY 743 PQDRLTEVIASIIKQNTYCVRRYAVVQKAAHGHVKAFAKSHVSTLTDLPYMRQFVAHL 802
Db 781 QETSPLRDAVVIQSSSLNEASSGLFDVFLRFMCHHAVIRKSVYVQCGIPQGSILSTL 840
QY 803 QETSPLRDAVVIQSSSLNEASSGLFDVFLRFMCHHAVIRKSVYVQCGIPQGSILSTL 862
Db 841 LCSLCYGDMMENKLFAGIRDDGLLLRLVDVDFLLVTPHLLTHAKTFLRTLVRGVPEYGCWNL 900
QY 863 LCSLCYGDMMENKLFAGIRDDGLLLRLVDVDFLLVTPHLLTHAKTFLRTLVRGVPEYGCWNL 922
Db 901 RKTVYNFPVEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVOSSDYSSYARTSIRASLTF 960
QY 923 RKTVYNFPVEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVOSSDYSSYARTSIRASLTF 982
Db 961 NRGFKAGRNMRKLFGLVRLKCHSLFLDLQVNSLQVTCVNIYKILLQAYRFHACVQLP 1020
QY 983 NRGFKAGRNMRKLFGLVRLKCHSLFLDLQVNSLQVTCVNIYKILLQAYRFHACVQLP 1042
Db 1021 FHQQVWKNTEFRLVSDTASLCYSTLKAKNAGMSLGAKAGPLPSEAVQWLCHQAFLL 1080
QY 1043 FHQQVWKNTEFRLVSDTASLCYSTLKAKNAGMSLGAKAGPLPSEAVQWLCHQAFLL 1102
Db 1081 KLTRHRVTYVPLLSLRTAQTLRSKLPGLTTLTALEAANPALPSDFKTILD 1132
QY 1103 KLTRHRVTYVPLLSLRTAQTLRSKLPGLTTLTALEAANPALPSDFKTILD 1154
RESULT 22
ID US-08-911-312-45 STANDARD; PRT; 1154 AA.
XX
AC xxxxxx
XX
XX
XX
DE Sequence 45, Application US/08911312
XX
XX Sequence 45, Application US/08911312
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase
CC NUMBER OF SEQUENCES: 170
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco

CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/026.981

INFO: 1132 AA; 126937 MW; 6
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1132 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE

CC TITLE OF INVENTION: Human Telomerase Catalytic Subunit
CC NUMBER OF SEQUENCES: 726
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US 08/974,549
CC FILING DATE: 19-NOV-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 23-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph, Red
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002610US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 344:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1132 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1132 AA; 126950 MW; 6598565 CN;
Query Match 98.0%; Score 8448; DB 18; Length 1132;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 1130; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 1 MPRAPRCRAVRLSHRYEVLPLATFVRRVLRGPGQWRLVQRGDPAAFRALVAQCLVCVPW 60
QY 23 MPRAPRCRAVRLSHRYEVLPLATFVRRVLRGPGQWRLVQRGDPAAFRALVAQCLVCVPW 82
Db 61 DARPPPAAPSFQVSCLEKELVARVLQRCERGAKNVLAFCGALLDGAAGGPPFAFTTSVR 120
|||||

QY 83 DARPPPAAPSFQVSCLEKELVARVLQRCERGAKNVLAFCGALLDGAAGGPPFAFTTSVR 142
Db 121 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLYLQGA 180
QY 143 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLYLQGA 202
Db 181 ATQARPPPHASGRRRLGRCERANNHVSREAGVPLGLPAPGARRRGGSSASRLPLKPRRR 240
QY 203 ATQARPPPHASGRRRLGRCERANNHVSREAGVPLGLPAPGARRRGGSSASRLPLKPRRR 262
Db 241 GAAPERTPVGQGSWAHPGRTGSPDRGFCVVSPPARPAEATSEALSGSTRHSHPVG 300
QY 263 GAAPERTPVGQGSWAHPGRTGSPDRGFCVVSPPARPAEATSEALSGSTRHSHPVG 322
Db 301 RQHAGPSTSRPPRMDTFCPPVYAEKHFYSSGDEQLRPSFLSSLRPSLTGARRL 360
QY 323 RQHAGPSTSRPPRMDTFCPPVYAEKHFYSSGDEQLRPSFLSSLRPSLTGARRL 382
Db 361 VETIFGSRWMPGCTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPRAAVT 420
QY 383 VETIFGSRWMPGCTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPRAAVT 442
Db 421 PAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSSPQVYGFVRACLRLVPPGLWGS 480
QY 443 PAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSSPQVYGFVRACLRLVPPGLWGS 502
Db 481 RHNERRLNRTKKFISLGKHAQLSLQELTWKMSVRDCAWLRSPGVCVPAAEHRLREEI 540
QY 503 RHNERRLNRTKKFISLGKHAQLSLQELTWKMSVRDCAWLRSPGVCVPAAEHRLREEI 562
Db 541 LAKFLHWMVYVVELLSRFFYVETTFQKNRLFYFPSPVSWKLSQSIGIRHKLKRVOLRE 600
QY 563 LAKFLHWMVYVVELLSRFFYVETTFQKNRLFYFPSPVSWKLSQSIGIRHKLKRVOLRE 622
Db 601 LSAEVRHREARPAALLTSRLRIPKDPGLRPIVNMDDYVVGARTFRERKRAERLTSRKA 660
QY 623 LSAEVRHREARPAALLTSRLRIPKDPGLRPIVNMDDYVVGARTFRERKRAERLTSRKA 682
Db 661 LFSVLYNERARRPGLLGASVLGLDDIHRWRTFVLRVRAQDPPPELYFVKVDVTGAYDTI 720
QY 683 LFSVLYNERARRPGLLGASVLGLDDIHRWRTFVLRVRAQDPPPELYFVKVDVTGAYDTI 742
Db 721 PQDLTEVIASIIKPNQTYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 780
QY 743 PQDLTEVIASIIKPNQTYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 802
Db 781 QETSPLEDAVVIQSSSLNEASSGLFDVLFREMFCHHVRIRGKSYVOCGIPGOSILSTL 840
QY 803 QETSPLEDAVVIQSSSLNEASSGLFDVLFREMFCHHVRIRGKSYVOCGIPGOSILSTL 862
Db 841 LCSLCYGD MENKLFAGIRRDGLLLRLVDDFLVTPHLLTHAKTFLRLTVRGVPEYGCYVNL 900
QY 863 LCSLCYGD MENKLFAGIRRDGLLLRLVDDFLVTPHLLTHAKTFLRLTVRGVPEYGCYVNL 922
Db 901 RKTVNFPEDEALGGTAFVQMPAHGLFPWCGLLLDPTLRTLEVSDSYSSYARTSTRSVTF 960
QY 923 RKTVNFPEDEALGGTAFVQMPAHGLFPWCGLLLDPTLRTLEVSDSYSSYARTSTRSVTF 982
Db 961 NRGFKAGRNMRRLFGVLRKCHSLFDLDVNSLQVCTNIYKILLQAVRFHACVLQLP 1020
QY 983 NRGFKAGRNMRRLFGVLRKCHSLFDLDVNSLQVCTNIYKILLQAVRFHACVLQLP 1042
Db 1021 FHQVWKNPTFFLRVISTASLCYSILKAKNAGMSLGAAGPLPSEAVOWLCHQAFLL 1080
QY 1043 FHQVWKNPTFFLRVISTASLCYSILKAKNAGMSLGAAGPLPSEAVOWLCHQAFLL 1102
Db 1081 KLTRHRYVYVPLGLSLRTAQTQSRKLPDGTTLTALEAAANPALPSDFKTLID 1132
QY 1103 KLTRHRYVYVPLGLSLRTAQTQSRKLPDGTTLTALEAAANPALPSDFKTLID 1154
RESULT 25
ID US-08-951-733-14 STANDARD: PRT: 949 AA.

MPERCH

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 24 08:00:10 1998; MasPar time 12.30 Seconds
Tabular output not generated. 664.418 Million cell updates/sec

Title: >US-08-951-733-20
Description: (1-1154) from US08951733.pep
Perfect Score: 8624
Sequence: 1 HASGRCVLLRTWELAPAT.....TALEAANPALPSDFKTLTD 1154

Scoring table: PAM 150
Gap 11

Searched: 77309 seqs, 7078906 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5_COMB 2:PCT9_COMB 3:backfiles1

Statistics: Mean 37.265; Variance 182.038; scale 0.205

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|------------|------------------------|-----------|
| 1 | 118 | 1.4 | 399 | 1 | US-08-147- | Sequence 27, Applicati | 2.06e+00 |
| 2 | 118 | 1.4 | 399 | 1 | US-08-447- | Sequence 27, Applicati | 2.06e+00 |
| 3 | 118 | 1.4 | 399 | 2 | PCT-US93-0 | Sequence 23, Applicati | 2.06e+00 |
| 4 | 118 | 1.4 | 399 | 1 | US-08-278- | Sequence 23, Applicati | 2.06e+00 |
| 5 | 118 | 1.4 | 399 | 1 | US-08-643- | Sequence 23, Applicati | 2.06e+00 |
| 6 | 118 | 1.4 | 399 | 1 | US-07-901- | Sequence 13, Applicati | 2.06e+00 |
| 7 | 118 | 1.4 | 399 | 1 | US-08-643- | Sequence 23, Applicati | 2.06e+00 |
| 8 | 118 | 1.4 | 399 | 1 | US-08-479- | Sequence 10, Applicati | 2.06e+00 |
| 9 | 118 | 1.4 | 399 | 1 | US-08-155- | Sequence 23, Applicati | 2.06e+00 |
| 10 | 118 | 1.4 | 399 | 2 | PCT-US93-0 | Sequence 23, Applicati | 2.06e+00 |
| 11 | 118 | 1.4 | 399 | 2 | PCT-US93-1 | Sequence 10, Applicati | 2.06e+00 |
| 12 | 118 | 1.4 | 399 | 2 | PCT-US93-0 | Sequence 23, Applicati | 2.06e+00 |
| 13 | 118 | 1.4 | 399 | 1 | US-08-462- | Sequence 23, Applicati | 2.06e+00 |
| 14 | 118 | 1.4 | 399 | 2 | PCT-US93-0 | Sequence 8, Applicatio | 2.06e+00 |
| 15 | 118 | 1.4 | 399 | 1 | US-08-451- | Sequence 23, Applicati | 2.06e+00 |
| 16 | 118 | 1.4 | 399 | 1 | US-08-480- | Sequence 23, Applicati | 2.06e+00 |
| 17 | 118 | 1.4 | 399 | 2 | PCT-US93-0 | Sequence 10, Applicati | 2.06e+00 |
| 18 | 118 | 1.4 | 399 | 2 | PCT-US93-0 | Sequence 23, Applicati | 2.06e+00 |
| 19 | 118 | 1.4 | 399 | 1 | US-08-406- | Sequence 23, Applicati | 2.06e+00 |
| 20 | 118 | 1.4 | 399 | 2 | PCT-US93-0 | Sequence 23, Applicati | 2.06e+00 |
| 21 | 118 | 1.4 | 399 | 2 | PCT-US93-0 | Sequence 23, Applicati | 2.06e+00 |
| 22 | 118 | 1.4 | 399 | 2 | PCT-US91-0 | Sequence 6, Applicatio | 2.06e+00 |
| 23 | 125 | 1.4 | 402 | 1 | US-07-841- | Sequence 29, Applicati | 7.05e-01 |

| | | | | | | | |
|----|-----|-----|-----|---|------------|------------------------|----------|
| 24 | 125 | 1.4 | 402 | 1 | US-08-643- | Sequence 21, Applicati | 7.05e-01 |
| 25 | 125 | 1.4 | 402 | 1 | US-08-451- | Sequence 21, Applicati | 7.05e-01 |
| 26 | 125 | 1.4 | 402 | 1 | US-08-452- | Sequence 21, Applicati | 7.05e-01 |
| 27 | 125 | 1.4 | 402 | 2 | PCT-US93-0 | Sequence 21, Applicati | 7.05e-01 |
| 28 | 125 | 1.4 | 402 | 1 | US-08-206- | Sequence 4, Applicatio | 7.05e-01 |
| 29 | 125 | 1.4 | 402 | 1 | US-08-155- | Sequence 21, Applicati | 7.05e-01 |
| 30 | 125 | 1.4 | 402 | 2 | PCT-US93-0 | Sequence 21, Applicati | 7.05e-01 |
| 31 | 125 | 1.4 | 402 | 1 | US-08-480- | Sequence 8, Applicatio | 7.05e-01 |
| 32 | 125 | 1.4 | 402 | 1 | US-07-901- | Sequence 11, Applicati | 7.05e-01 |
| 33 | 125 | 1.4 | 402 | 1 | US-08-147- | Sequence 29, Applicati | 7.05e-01 |
| 34 | 125 | 1.4 | 402 | 1 | US-08-643- | Sequence 21, Applicati | 7.05e-01 |
| 35 | 125 | 1.4 | 402 | 1 | US-08-406- | Sequence 21, Applicati | 7.05e-01 |
| 36 | 125 | 1.4 | 402 | 1 | US-08-479- | Sequence 8, Applicatio | 7.05e-01 |
| 37 | 125 | 1.4 | 402 | 1 | US-08-278- | Sequence 21, Applicati | 7.05e-01 |
| 38 | 125 | 1.4 | 402 | 1 | US-08-447- | Sequence 29, Applicati | 7.05e-01 |
| 39 | 125 | 1.4 | 402 | 2 | PCT-US93-0 | Sequence 21, Applicati | 7.05e-01 |
| 40 | 125 | 1.4 | 402 | 2 | PCT-US93-0 | Sequence 21, Applicati | 7.05e-01 |
| 41 | 125 | 1.4 | 402 | 2 | PCT-US93-0 | Sequence 21, Applicati | 7.05e-01 |
| 42 | 125 | 1.4 | 402 | 2 | PCT-US93-1 | Sequence 8, Applicatio | 7.05e-01 |
| 43 | 125 | 1.4 | 402 | 2 | PCT-US92-0 | Sequence 21, Applicati | 7.05e-01 |
| 44 | 125 | 1.4 | 402 | 2 | PCT-US93-0 | Sequence 11, Applicati | 7.05e-01 |
| 45 | 125 | 1.4 | 402 | 2 | PCT-US93-0 | Sequence 6, Applicatio | 7.05e-01 |

ALIGNMENTS

RESULT 1
ID US-08-147-023-27 STANDARD; PRT; 399 AA.
XX
AC xxxxxx
XX
DT

Sequence 27, Application US/08147023

Sequence 27, Application US/08147023

Patent No. 5468845

GENERAL INFORMATION:

APPLICANT: OPPERMAN, HERMANN

APPLICANT: OZKAYNAK, ENGIN

APPLICANT: KUBERASAMPATH, THANGAVEL

APPLICANT: RUEGER, DAVID C.

APPLICANT: PANG, ROY H. L.

TITLE OF INVENTION: OSTEOGENIC DEVICES

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ & THIBEAULT

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: U.S.A.

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/147,023

FILING DATE: 21-FEB-1992

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

PRIOR APPLICATION NUMBER: US 810,560

FILING DATE: 20-DEC-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 827,052

FILING DATE: 28-JAN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 660,162

FILING DATE: 22-FEB-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 621,988

FILING DATE: 04-DEC-1990

[illegible]

CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/155,343A
CC FILING DATE: 15-NOV-1993
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: FENTON ESQ., GILLIAN M.
CC REGISTRATION NUMBER: 36,508
CC REFERENCE/DOCKET NUMBER: CRP-067FW
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 248-7560
CC TELEFAX: (617) 248-7100
CC INFORMATION FOR SEQ ID NO: 23:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 399 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;

Query Match 1.4%; Score 118; DB 1; Length 399;
Best Local Similarity 38.7%; Pred. No. 2.06e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 GLALCALGGGHGPRPP-HTC-PQRLG-ARERDMOREILAVLGLPGR-PPRAQPAAR 67
QY 194 GPLYQLGAATQARPPPHASGPRRLGCRERAWNHVSREAGVPLGLPAPGARRRGSASRS 253

Db 68 QP 69
QY 254 LP 255

RESULT 10
ID PCT-US93-07231-23 STANDARD; PRT; 399 AA.
XX xxxxxx
DT
XX
DE Sequence 23, Application PC/TUS9307231
CC Sequence 23, Application PC/TUS9307231
CC GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: MORPHOGEN-INDUCED NERVE REGENERATION AND
CC TITLE OF INVENTION: REPAIR
CC NUMBER OF SEQUENCES: 33
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CREATIVE BIOMOLECULES, INC.
CC STREET: 35 SOUTH STREET
CC CITY: HOPKINTON
CC STATE: MASSACHUSETTS
CC COUNTRY: USA
CC ZIP: 01748
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/07231
CC FILING DATE: 19930729
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: KELLEY, ROBIN D.
CC REGISTRATION NUMBER: 34,637
CC REFERENCE/DOCKET NUMBER: CRP-070
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617/248-7000

CC TELEFAX: 617/248-7100
CC INFORMATION FOR SEQ ID NO: 23:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 399 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;

Query Match 1.4%; Score 118; DB 2; Length 399;
Best Local Similarity 38.7%; Pred. No. 2.06e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 GLALCALGGGHGPRPP-HTC-PQRLG-ARERDMOREILAVLGLPGR-PPRAQPAAR 67
QY 194 GPLYQLGAATQARPPPHASGPRRLGCRERAWNHVSREAGVPLGLPAPGARRRGSASRS 253

Db 68 QP 69
QY 254 LP 255

RESULT 11
ID PCT-US93-05446-13 STANDARD; PRT; 399 AA.
XX xxxxxx
DT
XX
DE Sequence 13, Application PC/TUS9305446
CC Sequence 13, Application PC/TUS9305446
CC GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: PROSTHETIC DEVICES HAVING ENHANCED
CC TITLE OF INVENTION: OSTEOGENIC PROPERTIES
CC NUMBER OF SEQUENCES: 22
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Creative BioMolecules, Inc.
CC STREET: 35 South Street
CC CITY: Hopkinton
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 01748
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/05446
CC FILING DATE: 19930608
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: PITCHER ESQ, EDMUND R
CC REGISTRATION NUMBER: 27,829
CC REFERENCE/DOCKET NUMBER: STR-057
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617/248-7000
CC INFORMATION FOR SEQ ID NO: 13:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 399 amino acids
CC TYPE: AMINO ACID
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;

Query Match 1.4%; Score 118; DB 2; Length 399;
Best Local Similarity 38.7%; Pred. No. 2.06e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 GLALCALGGGHGPRPP-HTC-PQRLG-ARERDMOREILAVLGLPGR-PPRAQPAAR 67
QY 194 GPLYQLGAATQARPPPHASGPRRLGCRERAWNHVSREAGVPLGLPAPGARRRGSASRS 253

QY 194 GPPLYQLGAATQARPPPHASGPRRLGCRAWNHSVREAGVPLGLPAPGARRRGGSSASRS 253
Db 68 QP 69
QY 254 LP 255

RESULT 12
ID PCT-US93-10520-10 STANDARD; PRT: 399 AA.
XX
AC xxxxxx
XX
DT
XX
DE
XX
Sequence 10, Application PC/TUS9310520
Sequence 10, Application PC/TUS9310520
GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: OP3-INDUCED MORPHOGENESIS
CC NUMBER OF SEQUENCES: 13
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CREATIVE BIOMOLECULES, INC.
CC STREET: 45 SOUTH STREET
CC CITY: HOPKINTON
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 01748
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION NUMBER: PCT/US93/10520
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/667,274
CC FILING DATE: 11-MAR-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/752,764
CC FILING DATE: 30-AUG-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/753,059
CC FILING DATE: 30-AUG-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/752,857
CC FILING DATE: 30-AUG-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/923,780
CC FILING DATE: 31-JUL-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/922,813
CC FILING DATE: 31-JUL-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: PITCHER ESQ, EDMUND R
CC REGISTRATION NUMBER: 27,829
CC REFERENCE/DOCKET NUMBER: CRP-076PC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (508)435-9001
CC INFORMATION FOR SEQ ID NO: 10:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 399 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;

Query Match 1.4%; Score 118; DB 2; Length 399;
Best Local Similarity 38.7%; Pred. No. 2.06e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

QY 194 GPPLYQLGAATQARPPPHASGPRRLGCRAWNHSVREAGVPLGLPAPGARRRGGSSASRS 253
Db 68 QP 69
QY 254 LP 255

RESULT 13
ID PCT-US93-07190-23 STANDARD; PRT: 399 AA.
XX
AC xxxxxx
XX
DT
XX
DE
XX
Sequence 23, Application PC/TUS9307190
Sequence 23, Application PC/TUS9307190
GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: MORPHOGEN-ENRICHED DIETARY COMPOSITION
CC NUMBER OF SEQUENCES: 33
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CREATIVE BIOMOLECULES, INC.
CC STREET: 35 SOUTH STREET
CC CITY: HOPKINTON
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 01748
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/07190
CC ATTORNEY/AGENT INFORMATION:
CC NAME: KELLEY, ROBIN D.
CC REGISTRATION NUMBER: 34,637
CC REFERENCE/DOCKET NUMBER: CRP-071
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617/248-7000
CC TELEFAX: 617/248-7100
CC INFORMATION FOR SEQ ID NO: 23:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 399 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;

Query Match 1.4%; Score 118; DB 2; Length 399;
Best Local Similarity 38.7%; Pred. No. 2.06e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

QY 194 GPPLYQLGAATQARPPPHASGPRRLGCRAWNHSVREAGVPLGLPAPGARRRGGSSASRS 253
Db 68 QP 69
QY 254 LP 255

RESULT 14
ID US-08-462-623-23 STANDARD; PRT: 399 AA.
XX
AC xxxxxx
XX
DT
XX
DE
Sequence 23, Application US/08462623

WILEY

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 08:54:27 1998; MasPar time 9229.89 Seconds
Tabular output not generated.
680.010 Million cell updates/sec

Title: >US-08-951-733-13
(1-2848) from US08951733.seq (1 of 3)
Perfect Score: 2848
N.A. Sequence: 1 CACGGCTCCGGCAGCGCTG.....GATGCCGCCACGCGCAT 2848
Comp: GTGCGCAGCGCGTGGCGAC.....CTACGGCGCGTGGCGGATA

Scoring table: TABLE default
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb155
1:em_ba 2:em_fun 3:em_htg 4:em_hun1 5:em_hun2 6:em_in
7:em_or 8:em_ov 9:em_pat 10:em_ph 11:em_pl 12:em_pl
13:em_ro 14:em_v1

Database: genbank107
15:gb_ba1 16:gb_ba2 17:gb_htg 18:gb_in 19:gb_cm 20:gb_ov
21:gb_pat 22:gb_ph 23:gb_pl1 24:gb_pl2 25:gb_pr1
26:gb_pr2 27:gb_pr3 28:gb_ro 29:gb_st 30:gb_sts 31:gb_sy
32:gb_un 33:gb_v1

Statistics: Mean 11.679; Variance 6.056; scale 1.928

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| SUMMARIES | | | |
|------------|-------|--------------------------|---|
| Result No. | Score | Query Match Length DB ID | Description Pred. No. |
| 1 | 2837 | 99.6 | 4015 27 AF015950 Homo sapiens telomeras 0.00e+00 |
| 2 | 2824 | 99.2 | 8960 31 AF043739 Synthetic construct hu 0.00e+00 |
| 3 | 2823 | 99.1 | 4027 27 AF018167 Homo sapiens telomeras 0.00e+00 |
| 4 | 825 | 29.0 | 3426 28 AF051911 Mus musculus telomeras 0.00e+00 |
| 5 | 823 | 28.9 | 3369 28 AF073311 Mus musculus telomeras 0.00e+00 |
| 6 | 129 | 4.5 | 201 28 AF029235 Mus musculus telomeras 1.63e-71 |
| 7 | 67 | 2.4 | 7218 21 I66494 Sequence 14 from paten 1.13e-25 |
| 8 | 49 | 1.7 | 7218 21 I66494 Sequence 14 from paten 1.72e-13 |
| 9 | 37 | 1.3 | 74371 26 AC005369 Homo sapiens chromosom 5.03e-06 |
| 10 | 34 | 1.2 | 215 21 I28278 Sequence 5 from patent 2.75e-04 |
| 11 | 35 | 1.2 | 215 21 I28278 Sequence 5 from patent 7.37e-05 |
| 12 | 34 | 1.2 | 74371 26 AC005369 Homo sapiens chromosom 2.75e-04 |

| | | | | | | |
|----|----|-----|--------|------------|------------------------|----------|
| 13 | 30 | 1.1 | 201 21 | A10158 | S.griseus gene for pre | 4.44e-02 |
| 14 | 30 | 1.1 | 201 21 | A10161 | Synthetic DNA for prep | 4.44e-02 |
| 15 | 30 | 1.1 | 201 21 | A10162 | Synthetic DNA for prep | 4.44e-02 |
| 16 | 30 | 1.1 | 201 21 | A10159 | S.griseus gene for pre | 4.44e-02 |
| 17 | 30 | 1.1 | 201 21 | E04076 | gDNA encoding envelope | 4.44e-02 |
| 18 | 31 | 1.1 | 216021 | HUAC004787 | Homo sapiens Chromosom | 1.28e-02 |
| 19 | 28 | 1.0 | 1663 | WU92534 | Mustela vison microsat | 4.95e-01 |
| 20 | 29 | 1.0 | 6011 | HS4RSIRZ2T | Epstein-Barr virus (wi | 1.50e-01 |
| 21 | 29 | 1.0 | 15807 | MEU72662 | Methylobacterium extor | 4.95e-01 |
| 22 | 28 | 1.0 | 31812 | SPUNK4 | S.pombe chromosome I c | 1.50e-01 |
| 23 | 28 | 1.0 | 39073 | SPAC2E11 | S.pombe chromosome I c | 4.95e-01 |
| 24 | 29 | 1.0 | 76094 | HS179D3A | Human DNA sequence fro | 1.50e-01 |
| 25 | 27 | 0.9 | 565 | E04076 | gDNA encoding envelope | 1.59e+00 |
| 26 | 27 | 0.9 | 904 | HS4BHLF | Epstein-Barr virus Bam | 1.59e+00 |
| 27 | 26 | 0.9 | 1341 | CHKD1LRB | Gallus domesticus DI-1 | 4.98e+00 |
| 28 | 26 | 0.9 | 2926 | SNAPROBA | S.marcescens proba ope | 4.98e+00 |
| 29 | 26 | 0.9 | 2926 | SMPROBAG | S.marcescens proba gen | 4.98e+00 |
| 30 | 26 | 0.9 | 4767 | AF034863 | Rattus norvegicus syna | 4.98e+00 |
| 31 | 27 | 0.9 | 5050 | HS4U2IR2A | Epstein-Barr virus (B9 | 1.59e+00 |
| 32 | 27 | 0.9 | 5400 | SCACTVA | S.coelicolor 6 actVA r | 1.59e+00 |
| 33 | 26 | 0.9 | 5591 | AF027499 | Pseudomonas aeruginosa | 4.98e+00 |
| 34 | 26 | 0.9 | 12949 | AF011922 | Azotobacter vinelandii | 4.98e+00 |
| 35 | 27 | 0.9 | 35980 | MTY20H10 | Mycobacterium tubercul | 1.59e+00 |
| 36 | 26 | 0.9 | 43034 | HSN5H5 | Human DNA sequence fro | 4.98e+00 |
| 37 | 27 | 0.9 | 64078 | AC004365 | Drosophila melanogaste | 1.59e+00 |
| 38 | 26 | 0.9 | 85132 | U82668 | Homo sapiens shox gene | 4.98e+00 |
| 39 | 26 | 0.9 | 123491 | HS126A5 | Human DNA sequence *** | 4.98e+00 |
| 40 | 26 | 0.9 | 125815 | HS436C18 | Human DNA sequence *** | 4.98e+00 |
| 41 | 26 | 0.9 | 138706 | AC004946 | *** SEQUENCING IN PROG | 4.98e+00 |
| 42 | 26 | 0.9 | 147216 | AF093117 | Homo sapiens chromosom | 4.98e+00 |
| 43 | 26 | 0.9 | 147542 | HS754E20 | Human DNA sequence *** | 4.98e+00 |
| 44 | 27 | 0.9 | 172281 | EBV | Epstein-Barr virus (EB | 1.59e+00 |
| 45 | 27 | 0.9 | 184113 | HS4B958RAJ | Epstein-Barr virus, ar | 1.59e+00 |

ALIGNMENTS

| | | | | | | |
|------------|---|----------|---------|------|-----|-------------|
| RESULT | 1 | AF015950 | 4015 bp | mRNA | PRI | 16-AUG-1997 |
| LOCUS | | | | | | |
| DEFINITION | | | | | | |
| ACCESSION | | | | | | |
| NID | | | | | | |
| KEYWORDS | | | | | | |
| SOURCE | | | | | | |
| ORGANISM | | | | | | |
| REFERENCE | | | | | | |
| AUTHORS | | | | | | |
| TITLE | | | | | | |
| JOURNAL | | | | | | |
| MEDLINE | | | | | | |
| REFERENCE | | | | | | |
| AUTHORS | | | | | | |
| TITLE | | | | | | |
| JOURNAL | | | | | | |
| FEATURES | | | | | | |
| source | | | | | | |
| gene | | | | | | |
| CDS | | | | | | |


```

1752 TGTACGGAGACCGTTCRAAAGACAGGCTCTTTTCTACCGAGAGAGTGCTCGAG 1811
1801 CAAGTTGCAAGCATTTGAATCAGACAGCATTTGAAGAGGTTGACGTGCGGAGCTGTC 1860
1812 CAAGTTGCAAGCATTTGAATCAGACAGCATTTGAAGAGGTTGACGTGCGGAGCTGTC 1871
1861 GGAAGCAGAGTTCAGCAGCATTCGGAGAGCCAGGCGCCCTGCTGAGCTCCAGACTCGG 1920
1872 GGAAGCAGAGTTCAGCAGCATTCGGAGAGCCAGGCGCCCTGCTGAGCTCCAGACTCGG 1931
1921 CTTTCATCCCAAGCTGAGCGGCTCGGCCGATTGTGAACATGGACTAGCTCGTGGGAGC 1980
1932 CTTTCATCCCAAGCTGAGCGGCTCGGCCGATTGTGAACATGGACTAGCTCGTGGGAGC 1991
1981 CAGAAGCTTCCAGAGAAAGAGGCGCGAGCGTCTCACTCGAGGGTGAAGCACTGTT 2040
1992 CAGAAGCTTCCAGAGAAAGAGGCGCGAGCGTCTCACTCGAGGGTGAAGCACTGTT 2051
2041 CAGCGTGTCAACTACGAGCGGCGCGGCCCGCCGCTCTGCGGCGCTCTGCTGCGG 2100
2052 CAGCGTGTCAACTACGAGCGGCGCGGCCCGCCGCTCTGCGGCGCTCTGCTGCGG 2111
2101 CCGTGCAGATATCCACAGGCGCTGGCGACCTTCGCTGCTGCTGCGGCGCCAGGACCC 2160
2112 CCGTGCAGATATCCACAGGCGCTGGCGACCTTCGCTGCTGCTGCGGCGCCAGGACCC 2171
2161 GCCGCTGAGCTGTACTTTGTCAAGTGTGATGAGCGGCGGTCAGACACCATCCCCCA 2220
2172 GCCGCTGAGCTGTACTTTGTCAAGTGTGATGAGCGGCGGTCAGACACCATCCCCCA 2231
2221 GGACAGGCTCAGGAGTTCATCGCCAGCATCATCAACCCAGACACAGTACTGCGTGG 2280
2232 GGACAGGCTCAGGAGTTCATCGCCAGCATCATCAACCCAGACACAGTACTGCGTGG 2291
2281 TCGGTATGCGGTGTCAGAGGCGCGCCATGCGGACGTCGCCAAGGCTTCAAGAGCA 2340
2292 TCGGTATGCGGTGTCAGAGGCGCGCCATGCGGACGTCGCCAAGGCTTCAAGAGCA 2351
2341 CGTCTCTACCTTGACAGACCTCCAGCGGTACATGCGACAGTTTCGTGGCTCACCTCAG 2400
2352 CGTCTCTACCTTGACAGACCTCCAGCGGTACATGCGACAGTTTCGTGGCTCACCTCAG 2411
2401 GACACGCGCTGAGGATGCGGTGTCATCGACGAGAGCTCTCTCCCTGAATGAGGCCAG 2460
2412 GACACGCGCTGAGGATGCGGTGTCATCGACGAGAGCTCTCTCCCTGAATGAGGCCAG 2471
2461 CAGTGGCTCTTCGAGCTCTCTTCAGCTTCATGTGCCACACGCGTGCATCAGGGG 2520
2472 CAGTGGCTCTTCGAGCTCTCTTCAGCTTCATGTGCCACACGCGTGCATCAGGGG 2531
2521 CAAGTCTCTACGTCCAGTCCAGGAGTCCCGAGGCTCCATCTCTCCACGCTGCTCTG 2580
2532 CAAGTCTCTACGTCCAGTCCAGGAGTCCCGAGGCTCCATCTCTCCACGCTGCTCTG 2591
2581 CAGCTGTGCTACGCGGACATCGAGACAAGCTGTTTGGGGGATTCGCGGGAGCGGCT 2640
2592 CAGCTGTGCTACGCGGACATCGAGACAAGCTGTTTGGGGGATTCGCGGGAGCGGCT 2651
2641 GCTCTCGCTTGTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCAGCGGAAC 2700
2652 GCTCTCGCTTGTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCAGCGGAAC 2711
2701 CTTCTCAGGACCTCGGTCGAGGTGTCCTCTGAGTATGCTCGGTGGTGAACCTCGGAA 2760
2712 CTTCTCAGGACCTCGGTCGAGGTGTCCTCTGAGTATGCTCGGTGGTGAACCTCGGAA 2771
2761 GACAGTGTGAATTCCTCTGAGACAGAGCGCTCGGTGGTGGCAGGCTTTGTTGACAT 2820
2772 GACAGTGTGAATTCCTCTGAGACAGAGCGCTCGGTGGTGGCAGGCTTTGTTGACAT 2831
2821 GCGGCGCCACGCGCTAT 2837
|||||
```

QY 2832 GCCGCCACGCGCTAT 2848

RESULT 2 AF043739 8960 bp DNA SYN 05-APR-1998
LOCUS
DEFINITION Synthetic construct human telomerase catalytic subunit
(hTERT/hEST2) mRNA and 3' Influenza A hemagglutinin tag, complete
sequence.
ACCESSION AF043739
NID 93023054
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
source
source
gene
CDS
source

synthetic construct.
synthetic construct
artificial sequence.
1 (bases 1 to 8960)
Telomerase activity is restored in human cells by ectopic
expression of hTERT (hEST2), the catalytic subunit of telomerase
Oncogene 16 (9), 1217-1222 (1998)
2 (bases 1 to 8960)
Counter, C.M., Meyerson, M., Eaton, E.N., Ellisen, L.W., Caddle, S.D.,
Haber, D.A. and Weinberg, R.A.
Direct Submission
Submitted (20-JAN-1998) Whitehead Institute, 9 Cambridge Center,
Cambridge, MA 02142, USA
Location/Qualifiers
1. 8960
/organism="synthetic construct"
/db_xref="taxon:32630"
order(1..1101,4606..8960)
/note="mammalian expression vector; see GenBank Accession
Number U47120"
/organism="Cloning vector pCI-neo"
/db_xref="taxon:45772"
1102..4542
/note="telomerase catalytic subunit hTERT/hEST2 mRNA"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5p15.33"
1102..4542
/gene="hTERT/hEST2"
1147..4605
/note="telomerase/reverse transcriptase"
/codon_start=1
/transl_table=11
/product="telomerase catalytic subunit"
/db_xref="PID:g3023055"
/translation="MPRAPRCRAVRSLLRSHRYEVLPLATFVRRLGPGQWRLVQGGDP
AAAFALVAQCLVCPMDAPPAPSPFQVSCLEKELVARVLQRLCERGANVLAFGFA
LLDARGGPPEAFTTSVRSYLPNTVDALGSGAWGLLRVGDVLLVHLLARALFV
LVAPGCAYQVCGPLXOLGAOTAPPPHAGSPRRLLGCSERANWHSYREAGVPLGLPA
PGARRGCSASRSLPLKPRRRCNAPEPTPVQGGSWAHGPRTRGSDRGFCVVSFA
RPAAEATSLGALSGTRHSPSVGRGHAGPPSTSRPPMDTPCPPVYATKFLYS
SGDEQLRPSFLSSRLTGAARLVETIFLGSRPWMPGTPRRLPLQRTWQMRVL
FLELLGHAQCPYGLLTKHCPFAAVTAAQVCAAREKPGCSVAAPPEEDTDRRLVQ
LLRHSPPQVQVYGFVRACLRLVPPGLWGRSHNRERFLRNTKFIISLGHAKLSLOEL
TWKSVRGCAWLRESQVCPAAEHRLREILAKFLHMLSVVVLRLSPFFVYVET
TFQKNLFFYKSKVWSKLQSIGIRHKLKRVQLRELSEAEVQREARPALITSLRIT
PKPDGLRPIVMDYVGARTFRREKRAERTSRVKALFVNLNERRRPPGLLGASVIG
LDDTHRAWRTFVLRQAQDPPPELYFKVDVTGAYDTIPQDRLTVEVASTIKPQNTYC
VRRYAVVQKAAGHVRKAFSHKSTLDLQAPYMQFVAHLQETSPLDADVIEQSSSL
NEASGLDFVLRPMCHAVIRKGSYVQCGIPQGSITSLTLCISLCYGDWENKLFAG
IRRDGLLLRLVDDELVTPLHTAKTFLRLVGRVPEYGCYVNLKTVNVPFVEDEAL
GGTAFVQMPAHGLFPGWGLLDDTTLTLEQSDYSYARTISIRASLTFRNGRAGNMR
KLFGVRLKCHSLFPLQVNSLQVCTNIYKILLQYRFRHACVLYLPFFHQVWKNRT
FFLRVISTASLCSYILKAKNAGSLGAKAGPLPSEAVQWLC HQFLKLTTRHYT
VPLGLSLRTAQTLQSLRKLPGTTLTALEAANPALPSDFKTLILDYPDVDPYAGPYD
4543..4605
/note="engineered based on the epitope for the


```
Qy 2001 CCGCAGAGAAAGAGGCGGCGCTCTCACCTCGAGGCTGAGGACACTGTTTCAGCGTCT 2060
Db 3141 CAACATACGAGCGGCGGCGGCGGCGCTCTCTGCGGCCCTCTGCTGCGCGCTGAGCA 3200
Qy 2061 CAACATACGAGCGGCGGCGGCGGCGGCGCTCTCTGCGGCCCTCTGCTGCGCGCTGAGCA 2120
Db 3201 TATCCACAGGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3260
Qy 2121 TATCCACAGGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2180
Db 3261 GCTGTACTTTTGTCAAGTGTGATGTGACGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3320
Qy 2181 GCTGTACTTTTGTCAAGTGTGATGTGACGGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2240
Db 3321 CACGAGGTCTATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3380
Qy 2241 CACGAGGTCTATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2300
Db 3381 CTTGTCTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3440
Qy 2301 CTTGTCTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2360
Db 3441 CTTGTCTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3500
Qy 2361 CTTGTCTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2420
Db 3501 GCTGAGGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3560
Qy 2421 GCTGAGGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2480
Db 3561 CTTGACAGCTCTTCTACGCTTCTATGTCGACACGCGGCGGCGGCGGCGGCGGCGGCGG 3620
Qy 2481 CTTGACAGCTCTTCTACGCTTCTATGTCGACACGCGGCGGCGGCGGCGGCGGCGGCG 2540
Db 3621 CTTGACAGCTCTTCTACGCTTCTATGTCGACACGCGGCGGCGGCGGCGGCGGCGGCG 3680
Qy 2541 CTTGACAGCTCTTCTACGCTTCTATGTCGACACGCGGCGGCGGCGGCGGCGGCGGCG 2600
Db 3681 CTTGACAGCTCTTCTACGCTTCTATGTCGACACGCGGCGGCGGCGGCGGCGGCGGCGG 3740
Qy 2601 CTTGACAGCTCTTCTACGCTTCTATGTCGACACGCGGCGGCGGCGGCGGCGGCGGCG 2660
Db 3741 TTTGCTGATGATTTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3800
Qy 2661 TTTGCTGATGATTTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3860
Db 3801 GACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3920
Qy 2721 GACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2780
Db 3861 GAACTTCCCTGTAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3920
Qy 2781 GAACTTCCCTGTAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2840
Db 3921 CGGCTAT 3928
Qy 2841 CGGCTAT 2848

RESULT 3
LOCUS AF018167 4027 bp mRNA PRI 29-AUG-1997
DEFINITION Homo sapiens telomerase catalytic subunit (hEST2) mRNA, complete
cde.
ACCESSION AF018167
NID 92347128
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 4027)
AUTHORS Meyerson,M., Counter,C.M., Eaton,E.N., Ellisen,L.W., Steiner,P.,
```

```
Caddle,S.D., Ziaugra,L., Beijersbergen,R.L., Davidoff,M.J., Liu,Q.,
Bacchetti,S., Haber,D.A. and Weinberg,R.A.
hEST2, the putative human telomerase catalytic subunit gene, is
up-regulated in tumor cells and during immortalization
Cell 90 (4), 785-795 (1997)
97433088
2 (bases 1 to 4027)
Meyerson,M., Counter,C.M., Eaton,E.N., Ellisen,L.W., Steiner,P.,
Caddle,S.D., Ziaugra,L., Beijersbergen,R.L., Davidoff,M.J., Liu,Q.,
Bacchetti,S., Haber,D.A. and Weinberg,R.A.
Direct Submission
Submitted (11-AUG-1997) Whitehead Institute for Biomedical
Research, Massachusetts Institute of Technology, Nine Cambridge
Center, Cambridge, MA 02142, USA
Location/Qualifiers
1. 4027
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5p15.33"
1. 4027
/notes="human Ever Shorter Telomeres 2"
/genes="hEST2"
/genes="hEST2"
/genes="hEST2"
/notes="member of reverse transcriptase class"
/codon_start=1
/db_xref="pid:g2347129"
/translation="MPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQNRLLVQRGDP
AARFALVAQCLVCPVDSRPPAPSPFRQVSKLKLVARVLQRCERKGNLAFGFA
LLDGARGGPEPFTTSPVSYLPTNTVDALRSGAGLLLRVDGDLVLRGALCALFV
LVAPSAQYQVGPPLYLQGAATARPAPPHASGPRRLGRCERANHSVREAGFVPLGPA
PGARRGGASRLPLPKRPRGAAPERTPVQGSWAHPRTGSRDGFVSPAP
PRAEATSLGALSGTRHSPVSGROHAGPSTSRPMDTSPVPAETKHELVS
SGKEQLRPSFLSSLLPSLTGARRLVETIFLGSRWMPGTPRRLPRLPORYQMRPL
FLELGNHAQCPYGLVLLKTLPLRAAVTPAAGVCAAREKPGSVAAPPEEDTPRRVQ
LLRQHSPPQVQYGFVRACLRLVPPGLWSRHRNERFLRNTKKFISLGRKALSLQEL
TKMSVSGCAWLRRSPGCVPAAEHLREELAKFLHLMVSVVLELLSFYVTF
TFQKRLFFYKSWKSLQSIGIRQHLKVLQRLSEAEVROHREARPLLTLRLRT
PKDGLRPIVMNDYVVGATGFRREKRAERLTSRKALFSVLYNERARPPGLIAGV
LDDIHRWRTFVLVRAQDPPPELYFVVDVTGAYDTIPDRLITEVTLASIKPONTYC
VRIAVVQKAAGHVRKAFKSHVSTLTDLPYMRQFVAHLQETSPLDVAIVBOSSL
NEASSGLDFVLFMCHHVRIRGKSVQCGQPGQISLTLCLSCYGDENKLFAL
IRDGILLRLVDDFLVLTPLHTAKTFLTLVRGVPYGCWNLKTVVNFVDEDEL
GTAFTVQMPAHLFPWCGLLDTLTLEVDSDYSYARTSIRASTLRNGFKAGNRMR
KLGVLRLKCHSLFLDLQVNSLQVCTNYKILLQAYREHACVLOLPPHQQVKNPT
FLRLVSDTASLICYSLKAKNAGMSLGAKAAGPLPSEAVQWLCVHQLKLRHRT
YVPLLSRLTAQTQLSRLKPLGTTTLTALEAANAALPSPDFKTLILD"
BASE COUNT 674 a 1361 c 1277 g 715 t
ORIGIN
Query Match 99.1%; Score 2823; DB 27; Length 4027;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 2825; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 12 GTCCTGCTGCGACGTGGGAAGCCCTGCGCCCGCCGACCCCGCGATGCCGCGGCTCCC 71
Qy 22 GTCTCTGTCGACAGTGGGAAGCCCTGCGCCCGCCGACCCCGCGATGCCGCGGCTCCC 81
Db 72 CGCTGCCGAGCGTGCCTCTCTGCGCAGCCACTACCGCAGAGGTGCTGCCGCTGCC 131
Qy 82 CGCTGCCGAGCGTGCCTCTCTGCGCAGCCACTACCGCAGAGGTGCTGCCGCTGCC 141
Db 132 ACCTTCGTCGCGGCGCTGGGGCGGCGGCGGCTGGCGGCTGGTGCAGCGCGGACCCGGCG 191
Qy 142 ACCTTCGTCGCGGCGCTGGGGCGGCGGCGGCTGGCGGCTGGTGCAGCGCGGACCCGGCG 201
Db 192 GCTTTCGCGCGCTGGTGGCGGCGGCGGCTGGTGGCGGCGGCGGCGGCGGCGGCGGCGG 251
Qy 202 GCTTTCGCGCGCTGGTGGCGGCGGCGGCGGCTGGTGGCGGCGGCGGCGGCGGCGGCGG 261
Db 252 CCGCGCGGCGGCGGCTGGTGGCGGCGGCGGCGGCTGGTGGCGGCGGCGGCGGCGGCGGCGG 311
```

|||||
QY 262 CCGCGGCCCTCTCTCCGCCAGGTGTCTGCTGGAAGAGCTGGTGGCCCCAGTGTG 321
Db 312 CAGAGCTGTGCGAGCGCGCGCGAAGACGTGCTGGCTTCGGCTTCGCGTGTGAG 371
QY 322 CAGAGCTGTGCGAGCGCGCGCGAAGACGTGCTGGCTTCGCGTGTGAG 381
Db 372 GGGGCGCGGGGGCCCCCGAGGCTTCACACACAGAGCGTGGAGCTACCTGCCAAC 431
QY 382 GGGGCGCGGGGGCCCCCGAGGCTTCACACACAGAGCGTGGAGCTACCTGCCAAC 441
Db 432 ACGGTGACCGACGCACTCGGGGGAGCGGGGTGGGGCTGCTGTGGCGCGGTGGG 491
QY 442 ACGGTGACCGACGCACTCGGGGGAGCGGGGTGGGGCTGCTGTGGCGCGGTGGG 501
Db 492 GACGAGCTGCTGTTACCTGTGCGACGCTGCGGCTCTTTGTGTGTGTGCTGCCAGC 551
QY 502 GACGAGCTGCTGTTACCTGTGCGACGCTGCGGCTCTTTGTGTGTGTGCTGCCAGC 561
Db 552 TGGGCTTACAGGTGTGGGGCGCGGCTGTACAGCTCGGCGCTGCCACTCAGGCGCGG 611
QY 562 TGGGCTTACAGGTGTGGGGCGCGGCTGTACAGCTCGGCGCTGCCACTCAGGCGCGG 621
Db 612 CCGCGCCACACGCTAGTGGACCCCGAAGCGCTCTGGGATGGAACGGGCTTGAACCAT 671
QY 622 CCGCGCCACACGCTAGTGGACCCCGAAGCGCTCTGGGATGGAACGGGCTTGAACCAT 681
Db 672 AGCGTCAGGAGCGCGGGTCCCTTGGGCTGCCAGCCCGGGGTGCGAGAGCGCGGG 731
QY 682 AGCGTCAGGAGCGCGGGTCCCTTGGGCTGCCAGCCCGGGGTGCGAGAGCGCGGG 741
Db 732 GCGAGTGCAGCGAAGTGTGCGGTTGCCAAGAGCGCCAGCGGTGGCGCTGCCCTGAG 791
QY 742 GCGAGTGCAGCGAAGTGTGCGGTTGCCAAGAGCGCCAGCGGTGGCGCTGCCCTGAG 801
Db 792 CCGGAGCGACCGCTTGGGCGAGGGTCTCGGGCCACCGCGGCGAGACGCTGGAGCG 851
QY 802 CCGGAGCGACCGCTTGGGCGAGGGTCTCGGGCCACCGCGGCGAGACGCTGGAGCG 861
Db 852 AGTACCGTGTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 911
QY 862 AGTACCGTGTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 921
Db 912 GAGGTGGCTCTGTGGACGCGGCACCTCCACCATCGTGGGCGCGAGCACCGCG 971
QY 922 GAGGTGGCTCTGTGGACGCGGCACCTCCACCATCGTGGGCGCGAGCACCGCG 981
Db 972 GSCCCCGCATCCACATCGCGGCGCACCGTCCCTGGGACAGCGCTTGTCCCGCGGTGAC 1031
QY 982 GSCCCCGCATCCACATCGCGGCGCACCGTCCCTGGGACAGCGCTTGTCCCGCGGTGAC 1041
Db 1032 GCCGAGACCAAGCACTTCTCTACTCTCTAGCGCACAGGAGCACTGCGGCCCTCTTC 1091
QY 1042 GCCGAGACCAAGCACTTCTCTACTCTCTAGCGCACAGGAGCACTGCGGCCCTCTTC 1101
Db 1092 CTACTAGCTCTGTAGGCGCGAGCTGACTGCGGCTCGGAGCTCGTGGAGCACTTTT 1151
QY 1102 CTACTAGCTCTGTAGGCGCGAGCTGACTGCGGCTCGGAGCTCGTGGAGCACTTTT 1161
Db 1152 CTGGGTTCAGGCGCTTGGATCGAGGACTCCCGCAGGTTGCCCGCGCTGCCAGCGC 1211
QY 1162 CTGGGTTCAGGCGCTTGGATCGAGGACTCCCGCAGGTTGCCCGCGCTGCCAGCGC 1221
Db 1212 TACTGGCAATCGGCGCTGTTCTTGGAGCTGTTGGGAACACAGCGAGTGGCCCTAC 1271
QY 1222 TACTGGCAATCGGCGCTGTTCTTGGAGCTGTTGGGAACACAGCGAGTGGCCCTAC 1281
Db 1272 GGGGTGCTCTCAAGACGACTGCGCGTGTGAGTGTGGGTCAACCCAGAGCGGCTGTC 1331
QY 1282 GGGGTGCTCTCAAGACGACTGCGCGTGTGAGTGTGGGTCAACCCAGAGCGGCTGTC 1341
Db 1332 TGTGCGCGGAGACCGCGAGGCTGTGTGGCGCGCGCGGAGGAGGAGACACACCGC 1391
|||||

QY 1342 TGTGCCCGGGAAGACCCCAAGGGCTGTGTGGCGCGCGCGGAGGAGGACACAGACCC 1401
Db 1392 CGTGCCTGTGTGACAGCTGCTCCGCCAGCACAGACGCGCTTGGAGGTGTAGGGTTCGTG 1451
QY 1402 CGTGCCTGTGTGACAGCTGCTCCGCCAGCACAGACGCGCTTGGAGGTGTAGGGTTCGTG 1461
Db 1452 CGGGCTGCTGTGCGCGGCTGTGCGGCCAGGCTCTGGGGCTCCAGGACCAAGAACGC 1511
QY 1462 CGGGCTGCTGTGCGCGGCTGTGCGGCCAGGCTCTGGGGCTCCAGGACCAAGAACGC 1521
Db 1512 CGCTTCCTCAGGAACACCAAGATTCTATCTCCCTGGGGAAGCATGCCAAGCTTCGCTG 1571
QY 1522 CGCTTCCTCAGGAACACCAAGATTCTATCTCCCTGGGGAAGCATGCCAAGCTTCGCTG 1581
Db 1572 CAGGAGCTGAGCTGGAAGATGAGGCTGCGGGCTGCGCTTGGCTGCGAGAGCCAGGG 1631
QY 1582 CAGGAGCTGAGCTGGAAGATGAGGCTGCGGGCTGCGCTTGGCTGCGAGAGCCAGGG 1641
Db 1632 GTTGGCTGTGTTCCGGCGCGAGACACGCTGTGCTGTGAGGAGATCTTGGCCAAATTCCTG 1691
QY 1642 GTTGGCTGTGTTCCGGCGCGAGACACGCTGTGCTGTGAGGAGATCTTGGCCAAATTCCTG 1701
Db 1692 CACTGGCTGTAGTGTGTACGTGCTGAGCTGCTCAGGCTCTTTCTTTATGTACCGAG 1751
QY 1702 CACTGGCTGTAGTGTGTACGTGCTGAGCTGCTCAGGCTCTTTCTTTATGTACCGAG 1761
Db 1752 ACCAGCTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCA 1811
QY 1762 ACCAGCTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCA 1821
Db 1812 AGCATTTGAATCAGACAGCACTTGAAGAGGTGCAGTGTGCGGAGCTGTCTGGAGAGCAG 1871
QY 1822 AGCATTTGAATCAGACAGCACTTGAAGAGGTGCAGTGTGCGGAGCTGTCTGGAGAGCAG 1881
Db 1872 GTCAGGAGCATCGGAAGCCAGCCCGCTGCTGAGTCCAGCTCCGCTTATCCCC 1931
QY 1882 GTCAGGAGCATCGGAAGCCAGCCCGCTGCTGAGTCCAGCTCCGCTTATCCCC 1941
Db 1932 AAGCTGACGGGCTGCGGCGGCTTGTGAACATGGACTACGTCGTGGGAGCCAGAACGTT 1991
QY 1942 AAGCTGACGGGCTGCGGCGGCTTGTGAACATGGACTACGTCGTGGGAGCCAGAACGTT 2001
Db 1992 CGCAGAGAAAGAGGGCGGAGCGTCTCAGTCGAGGGTGAAGGCACTGTTTCAGCGTGTGTC 2051
QY 2002 CGCAGAGAAAGAGGGCGGAGCGTCTCAGTCGAGGGTGAAGGCACTGTTTCAGCGTGTGTC 2061
Db 2052 AACTACAGGCGGCGCGCGCGCTCTCTGCGCGCTCTGCTGCGGCTTGAGCGAT 2111
QY 2062 AACTACAGGCGGCGCGCGCGCTCTCTGCGCGCTCTGCTGCGGCTTGAGCGAT 2121
Db 2112 ATCCACAGGGCTGCGCACCTTCTGCTGCTGTGGGGCCCGAGGACCCCGCGCTGAG 2171
QY 2122 ATCCACAGGGCTGCGCACCTTCTGCTGCTGTGGGGCCCGAGGACCCCGCGCTGAG 2181
Db 2172 CTGTACTTTTCAAGGTGGATGTACGCGGCGGTACGACACCATCCCCCAGGACAGGCTC 2231
QY 2182 CTGTACTTTTCAAGGTGGATGTACGCGGCGGTACGACACCATCCCCCAGGACAGGCTC 2241
Db 2232 ACGGAGTCTACGCCAGCATCAAAACCCAGAACAGTACTGCTGCTGCTGATGCTG 2291
QY 2242 ACGGAGTCTACGCCAGCATCAAAACCCAGAACAGTACTGCTGCTGCTGATGCTG 2301
Db 2292 GTGTCTCAGAGGGCGGCCATGGCAGCTCGCAAGGCTTCAAGAGCCAGCTCTCTACC 2351
QY 2302 GTGTCTCAGAGGGCGGCCATGGCAGCTCGCAAGGCTTCAAGAGCCAGCTCTCTACC 2361
Db 2352 TTGACAGACCTTCCAGCGGTACATGCGAGTTCGTGGCTACCTTGCAGGAGACAGGCCG 2411
QY 2362 TTGACAGACCTTCCAGCGGTACATGCGAGTTCGTGGCTACCTTGCAGGAGACAGGCCG 2421
Db 2412 CTGAGGAGTCCGCTGCTATCGAGCAGAGCTCTCTGCTGATGAGGAGCAGTGGGCTC 2471
QY 2422 CTGAGGAGTCCGCTGCTATCGAGCAGAGCTCTCTGCTGATGAGGAGCAGTGGGCTC 2481

```
Db 2472 TTCAGCTCTTCTACGCTTCATGTGCCACACGCGCGTGGCATCAGGGCGCAAGTCTCTAC 2531
|||||
QY 2482 TTCAGCGTCTTCTACGCTTCATGTGCCACACGCGCGTGGCATCAGGGCGCAAGTCTCTAC 2541
|||||
Db 2532 GTCAGTGCCAGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTCAGCGCTGTGC 2591
|||||
QY 2542 GTCAGTGCCAGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTCAGCGCTGTGC 2601
|||||
Db 2592 TACGGCGACATGAGAACAGCTGTTTGGCGGGATTGCGCGGACGGGCTGCTCTCGCT 2651
|||||
QY 2602 TACGGCGACATGAGAACAGCTGTTTGGCGGGATTGCGCGGACGGGCTGCTCTCGCT 2661
|||||
Db 2652 TTGGTGATGATTCTTGTGTGACACCTACCTACCTACCTACCTACCTACCTACCTACCT 2711
|||||
QY 2662 TTGGTGATGATTCTTGTGTGACACCTACCTACCTACCTACCTACCTACCTACCTACCT 2721
|||||
Db 2712 ACCTGTGCGAGGTGCTCTGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2771
|||||
QY 2722 ACCTGTGCGAGGTGCTCTGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2781
|||||
Db 2772 AACTTCCCTCTAGAACAGAGGCGCTGGTGGCAGCGCTTTTGTTCAGATGCGCGGCCAC 2831
|||||
QY 2782 AACTTCCCTCTAGAACAGAGGCGCTGGTGGCAGCGCTTTTGTTCAGATGCGCGGCCAC 2841
|||||
Db 2832 GGCCTAT 2838
|||||
QY 2842 GGCCTAT 2848
|||||

RESULT 4
LOCUS AF051911 3426 bp mRNA ROD 02-APR-1998
DEFINITION Mus musculus telomerase reverse transcriptase mRNA, complete cds.
ACCESSION AF051911
NID 93005591
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 3426)
AUTHORS Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
TITLE Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.
REFERENCE 2 (bases 1 to 3426)
AUTHORS Expression of mouse telomerase reverse transcriptase during
TITLE development, differentiation, and proliferation
JOURNAL Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.
SUBMITTED (02-MAR-1998) Microbiology and Immunology, Albert
Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY
10461, USA
FEATURES
source Location/Qualifiers
1..3426
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="13"
/map="MGD offset position 40; 40 cm from centromere"
30..3398
/note="mTERT"
/codon_start=1
/product="telomerase reverse transcriptase"
/db_xref="PID:g3005592"
/translation="MTRAPRCAPVRSLLRSRYREVWPLATFVRRLGPGRRLLVQPGDP
KIVTLVAOCLVCMHWGSPPPADLSHFQVSSILKELVAVRVORLERNERNVLTGFE
LNEARGPMAFTSVRSYLVNTVETLRVSGANWMLLSRVGDDLLVYLALCALYL
LVPSCAYQVCSPLQICATDIMPVSASRYRTPRVGNFTNDRFLQIQQSSSSRQE
APKPLALSRGTRKHLSLTSTSVPSAKKARCPVPRVEEGPHRVLPTEGSKSWPSP
ARSPYPTAEKDLSSKRVSDLSGVCCKHPKSPSTLSLPPRQNAFOLRPFIEFRH
FLYSRGDGERLNPFLNLOPLNLTGARLVEIIFLGRPRPTSGPLCRLHLSRRYV
QMRPLFOQLLVNHAECQVYVRLLSRHSCHREFTANQOVDALNTSPPHLMDLLRLHSSPWQ
VYGLFACILCKVYVSLWGTNRNRRFFNLKFKFISLGYKLSLQELMWMKVEDCH
WLKSPGKORVPAAEHRLEIRILATFLWMDTYVVOILLRSFFYITESTFQKNRLFFY
RKSWNSKLSIGVROHLERVRLRELSQEEVRRHQDTWLAIPCRLRFLIPKPNGLRPIV
```

```
NMSYSGTRALRRKQAOHFTORLKTLEFSLMNYERTKHPHLMGSSVLGMNDIYRTWRA
FVLRVALDQTPRMFYRADVTGAYDAIPQGLVYEVANMIRHSSTCYCIRYAVVRR
DSQGVHKSFRQVTTLSLQPYMGQFLKHLQSDASALRNSWVTEQISNMNESSSL
FDFFLHFRHSVVKIGRCYTCQCIPOGSSSLTCLSCFDMENKLFVQVROGLL
LRFVDDFLVTPHLDQAKTFLSTLVHGVPEYGMINLOKTVVNEPVEPTGLGAAPYQ
LPAHCLFPKCGILLDTQLEVEFCDYSGVAQTSIKTSLTFSQVFKAGKMRNKLVLRL
LKCHGLFDLQVNSLQTVCIINIYKIFLQOARFHCYIQLPFDPRKRLNLTFLGITS
SQASCCYAILKVNPMGTMKASGSPFPAHWCYQAFLLKLAHSHVIYKLCULGPLRT
AOKLCLRKLPENTMTILKAAADPALSTFQITLD"
BASE COUNT 746 a 979 c 907 g 794 t
ORIGIN
Query Match 29.0%; Score 825; DB 28; Length 3426;
Best Local Similarity 67.1%; Pred. No. 0.00e+00;
Matches 1901; Conservative 0; Mismatches 838; Indels 93; Gaps 5;
Db 14 CCGGGCTTTGAGCACAAATGACCGCGCTCTCGTTCCCGCGGTGCGTCTCTGCTGCG 73
|||||
QY 51 CCGGGCACCCCGCGATGCGGGCGCTCCCGCTCCGAGCGCTGCTCCCTGCTGCG 110
|||||
Db 74 CAGCCGATACCGGAGGTGTGCGCGCTGGCAACCTTTGTGCGCGCTGGGCGCCGAGGG 133
|||||
QY 111 CAGCCACTACCGGAGGTGTGCGCGCTGGCAACCTTTGTGCGCGCTGGGCGCCGAGGG 170
|||||
Db 134 CAGGCGCTTGTCAACCGCGGACCCGAGATCTACCGCACTTTGTTGCCCAATGCCT 193
|||||
QY 171 CTGGCGCTGTGTGACGCGGCGACCCGCGCTTTTCGCGCGCTGTGCGCCAGTGCCT 230
|||||
Db 194 AGTGTGATGCTGCTGCGGCTTCACAGCTCCACCTGCGGACCTTTCTCCACAGTGTG 253
|||||
QY 231 GGTGTGCGTCCCTGGGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 290
|||||
Db 254 ATCCCTGAAAGAGCTGGTGCCAGGCTTGTGACAGACTTGTGAGCGCAACGAGAGAA 313
|||||
QY 291 CTGCTGAGAGAGCTGGTGCGCGAGTGTGACAGAGCTGTGCGAGCGCGCGCGAGAA 350
|||||
Db 314 CTGCTGCGCTTTTGGCTTTGAGCTGTTAACGAGGCGAGAGCGCGGCTTCCATGCGCTT 373
|||||
QY 351 CGTCTGCGCTTTCGCTTTCGCTGCTGACGCGGCGCGCGCGCGCGCGCGCGCGCTT 410
|||||
Db 374 CACTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 433
|||||
QY 411 CACACAGCGCTGGGAGCTTACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 470
|||||
Db 434 TGCTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 493
|||||
QY 471 GCGTGGGCGCTGCTGCTGCGCGCGCTGGGCGAGCTGCTGCTGCTGCTGCTGCTGCTG 530
|||||
Db 494 CTGCTGCTTTTATCTTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 553
|||||
QY 531 CTGCGCGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 590
|||||
Db 554 GTACCAAAATTTGGTGGCACACGAGATATCTGGCGCTCTGTGCTGCTGCTGCTGCTGCTG 613
|||||
QY 591 GTACCAAGCTGGGCGCTGCGCACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 650
|||||
Db 614 CCGACCGCTGGGCGAGGAATTTCACTAACCTTAGTGTCTTACAACAGATCAAGAGCAGTAG 673
|||||
QY 651 GCCTCTGGGATGCG-----AACGGCGCTGGAAACATAG 693
|||||
Db 674 TCGCCAGGAGACACCGAAACCCCTGGCTTGCATCTCGAGGTACAAAGAGGATCTGAG 733
|||||
QY 684 CGTCAGGAGCGCGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743
|||||
Db 734 TCTACCACTAAGTGTGCTTCACTTACAGTACAGGCGCAGATGCTATCTCTGCTCCGAGAGT 793
|||||
QY 744 CAGTGGCAGCGGAAGTGTGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 803
|||||
Db 794 GGAGGAGGACCCACACGCGAGGTGCTACCAACCCCATCAGGCAAAATCATGGTGTGCAAG 853
|||||
QY 804 GGAGCGGACCGCGCTGGGCGAGGGGTCTCTGGGCGCGCGCGCGCGCGCGCGCGCGCGAG 863
|||||
Db 854 TCCTGCTGCTCCCGCGAGGTGCTTACTGACAGAAAGATTTGCTCTTAAAGAGAAAGGT 913
|||||
```


| REFERENCE | 1 | (bases 1 to 3369) |
|-----------------------|---|-------------------|
| AUTHORS | Martin-Rivera, L., Herrera, E., Albar, J.P. and Blasco, M.A. | |
| TITLE | Expression of mouse telomerase catalytic subunit in embryos and adult tissues | |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 95 (18), 10471-10476 (1998) | |
| MEDLINE | 98393668 | |
| REFERENCE | 2 | (bases 1 to 3369) |
| AUTHORS | Martin-Rivera, L., Herrera, E. and Blasco, M.A. | |
| TITLE | Direct Submission | |
| JOURNAL | Submitted (19-JUN-1998) Immunology and Oncology, National Centre of Biotechnology, Cantoblanco, Madrid 28049, Spain | |
| FEATURES | Location/Qualifiers | |
| source | 1. .3369 | |
| | /organism="Mus musculus" | |
| | /db_xref="taxon:10090" | |
| CDS | 1. .3369 | |
| | /codon_start=1 | |
| | /product="telomerase catalytic subunit" | |
| | /db_xref="PDB:g3551847" | |
| | /translation="MTRAPRCPAVRLLRSRYREVWPLATFVRRLLGPEGRRLVQPGDP KIYRTVAOCLVCHMGHSOPPPADLTFHQVLSKELMARVVRVQCLCRNERNVLAFIGE LNEARGCCPMATFTSVRSYVLTPIETVTRVSGAMLLPNTVSRVGDLLVYLALCALVL LVPPSACVQVCGSLQICATTDIPWSVASRYPTRPVGENFTNLRFLQIKSSROE APRPLALPSGTGRHLISLSTSVPSAKKARCYPPRVVEGPHQVLPVTPSGKSWPFS ARSPVETAKDLSVGVCKKHPKPSSTLSLPPRQNAFLQRPETFRH FLYRSDGDEQVPEPFLNLLNPNTGARLVEIIFLGSRPSTSGLCRTHLRSRYR QMPRLQQLLVNHAQYVLLRSHCFRTANQOQTDALNTSPHMLDLRLKHSVQW VYGFRLACILKVVSASHLRHRERFFKNLKFISLQYKGLSLQELAMKMKVEDCH WLRSPKORVPAEHLRERILATFLWMDTVVQLLSRFFYTESFTQKRLFFV RKSVSKLSIGVQHLRVRVRLSELQSEVRRHQDTWLAMPICRLRPIKPNGLRPIV NMSYSGMTALRGARKQAQHTQKTLTFLSMYERTKHPHMGSSVLGMNDIYRTVRA FVLRLADQTPMYFVKADYQDAIPQGLVEYVANNIRHSESTYCIQYAVVRV DSQGVHKSFRQVITLSDLPQWGFQKHLQSDASALRNSVVIQSIISMNSDSL DFDLHFRSVVKIGDRCTQCGIPOGSSLSLTLCLSCFCGDMENKLPFAEYVORDGL LRVDDFLVTPHQDKATFLSLVHGVEPYGCMINLQTVNFPVEPTLGGAAAPYQ LPACHLFPWCGLLDTQTLVECDYGIQATSIKTLSTFQSVFAKATMRNKLKLSVLR LKCHQVLQVNSLTQVINYIKYFLQAYRHACVIOLPDQVRNRNNTFFLGIIS SQAQSCYAILKVNQMTLKASGTFPPEAHWLCYQAFLLKLAHSVIYRCKLLGLPTL AOKRLKPEATMTILKAADPALSTDFQITLD" | |
| BASE COUNT | 737 a 959 c 891 g 782 t | |
| ORIGIN | | |
| Query Match | 28.9%; Score 823; DB 28; Length 3369; | |
| Best Local Similarity | 67.2%; Pred. No. 0.00e+00; | |
| Matches | 1892; Conservative 0; Mismatches 831; Indels 93; Gaps 5; | |
| Db | 1 ATGACCCGCGCTCTCGTGGCCCGCGGTGGCTCTCTGCTGGCAGCGCATACGGGAG 60 | |
| Qy | 1 ATGCGCGCGCTCCCCGTCGCGAGCGCTGGCTCCCTGCTGGCAGCCTACCGCGAG 126 | |
| Db | 61 GTGTGGCGCTGGCAACCTTTGTGCGGCGCTGGGGCCGAGGCGAGCGGCTTTGTGCAA 120 | |
| Qy | 127 GTGTGGCGCTGGCCACGTTGCTGCGGCGCTGGGGCCGAGGCGTGGCGCTGTGCGAG 186 | |
| Db | 121 CCGGGGACCCGGAAGATCTACCGCACTTTGGTTGCCCAATGCCCTAGTGTGCATCTGG 180 | |
| Qy | 187 CCGGGGACCCGGGGGCTTTCCCGCGCTGTGTGCCCAAGTGCCTGGTGGCTGCGCTGG 246 | |
| Db | 181 GGCTCACAGCTCCACCTGGCGACCTTCCCTTCACACAGGTGTGTCATCCCTGAAGAGCTG 240 | |
| Qy | 247 GACGACGGCGCGCGCGCGCCCTCCCTTCCTCCGCGAGGTGTCTCTGCTGAAGAGCTG 306 | |
| Db | 241 GTGGCCAGGGTTGTGCAGAGACTCTGCGAGCGCAACAGAGAAAGCTGCTGGCTTTTGGC 300 | |
| Qy | 307 GTGGCCGAGTGTGTGACAGAGCTGTGCGAGCGCGGCGGAAGACGTGCTGGCTTGGC 366 | |
| Db | 301 TTTGAGCTGCTTAACGAGGCCAGAGCGGGGCTCCCATGGCTTCACTAGTAGGCTGGCT 360 | |
| Qy | 367 TTCGCGCTGTGGACGGGGCGCGGGGGGCCCCCGAGGCTTCCACACAGCGTGGC 426 | |
| Db | 361 AGCTACTTGGCCACACTGTTATTAGACACCTCGGTGTGCTAGTGTGCATGATGCTACTG 420 | |
| Qy | 427 AGCTACTTGGCCACAGGTCACGAGCACTGGGGGAGCGGGCGCTGGGGGCTGCTGCTG 486 | |

```
Db 1444 AAGAGTTTCATCTCGTTGGGAATACGGCAAGCTATCACTCCAGGAAGTATGATGTGGAAG 1503
|||||
Qy 1540 AAGAGTTTCATCTCGTTGGGAAGATCCCAAGCTCTCGTGCAGGAGCTGACGTGGGAAG 1599
|||||
Db 1504 ATGAAGTAGAGGATTCACCTGGCTCGCGACAGCCCGGGGAAGGACCGTGTCCCGGCT 1563
|||||
Qy 1600 ATGAGCGTGGGAGTGGCTTGGCTGGCGAGGAGCCAGGGGTGGCTGTGTCCCGGCC 1659
|||||
Db 1564 GCAGAGCACCGTCTGAGGAGGAGGATCTGCTAGTTCCTGTCTGCTGATGGACACA 1623
|||||
Qy 1660 GCAGAGCACCGTCTGCTGAGGAGATCTGGCCAAAGTTCTCTGCACTGGCTGATGAGTGTG 1719
|||||
Db 1624 TACGTGGTACAGCTGCTTAGTTCATCTTTTACATCAGAGAGGACATCCAGAGAAC 1693
|||||
Qy 1720 TACGTGGTACAGCTGCTCAGTCTCTTTTATGTACGGAGCCACGCTTTCAAAGAAC 1779
|||||
Db 1684 AGGCTCTTCTTACCGTAAGAGTGTGGGAGCAAGCTGCAGAGCATGGAGTCAGGCAA 1743
|||||
Qy 1780 AGGCTCTTCTTACCGTAAGAGTGTGGGAGCAAGTGTGAGCAAGTGTCAAGCATGGAAATCAGACAG 1839
|||||
Db 1744 CACCTTGAGAGAGTGGCTACGGGAGCTGTCAACAAGAGGAGGTCAAGCATCACAGGAC 1803
|||||
Qy 1840 CACTTGAAGAGGTGCAGCTCGGAGCTGTGGAAGCAGAGGTCAAGCATCGGGAA 1899
|||||
Db 1804 ACCTGGCTAGCATGCCATCTGCAGACTGCGTTCATCCCAAGCCCAACGGGCTGGCG 1863
|||||
Qy 1900 GCCAGGCCGCGCTGCTGACCTCCAGACTCCGCTTCATCCCAAGCCCTGACGGCTGGCG 1959
|||||
Db 1864 CCCATTGTGAACATGAGTTATAGATGGGTACCAAGAGCTTTGGCAGCAAGCAAGCAGGCC 1923
|||||
Qy 1960 CCGATTGTGAACATGAGTACGTCTGGGAGCCAGAGCTTCCGAGAGAAAGAGGGCC 2019
|||||
Db 1924 CAGCATTTCCACCGGCTCTCAAGACTCTCTTCAGCATGCTCAACTATGAGCGCAAAA 1983
|||||
Qy 2020 GAGCGTCTCACTCGAGGCTGAAGCACTGTTACGCGTGTCTCAACTAGCGCGCGCGG 2079
|||||
Db 1984 CATCTCACTTATGGGCTCTCTGTACTGGGTATGATGATCATCTACAGCACTGGCGG 2043
|||||
Qy 2080 GCGCCGCGGCTCTCTGGGCGCTCTGTGCTGGCGCTGGAGATCCACAGGCGCTGGCGC 2139
|||||
Db 2044 GCCTTTGTGCTGCGTGTGCTGTGAGCAGACACCCAGGATGTACTTTGTTAAGCA 2103
|||||
Qy 2140 ACCTTGTGCTGCGTGTGCGGCGCCAGGACCCGCGCTGAGCTGTACTTTGTCAAGGTG 2199
|||||
Db 2104 GATGTACCGGGGCTATGATGCCATCCCGAGGCTAAGCTGGTGGTGGTGTGTTGCCAAT 2163
|||||
Qy 2200 GATGTACGGGCGGTACGACACCATCCCGCAGGACAGGCTCACGGAGGTCTATCGCCAGC 2259
|||||
Db 2164 ATGATCAGGCACTCGGAGAGCAGCTACTGTATCCGCCAGTATGAGTGGTCCGGAGAGAT 2223
|||||
Qy 2260 ATCATCAACCC---AGAACACGTACTGCGTGTGCGTATGCGGTGGTCCAGAGGCC 2316
|||||
Db 2224 AGCCAAGGCCAAGTCCACAAGTCTTTAGGAGACAGAGGTACACCCCTCTCTGACCTCCAG 2283
|||||
Qy 2317 GCCATGGGACGTCCGCAAGGCTTCAAGAGCCACGTCTCTACTTTCAGAGACCTCCAG 2376
|||||
Db 2284 CCATACATGGGCCAGTTCCTTAAGCATCTGAGGATTCAGATGCCAGTGCACACTGAGGAAC 2343
|||||
Qy 2377 CCGTACATGCGACAGTTCGTGGCTCACCTGCAGGAGAC-----CAGCCCGCTGAGGAT 2430
|||||
Db 2344 TCCGTTGTATCAGCAGAGCATCTATGATGAGAGCAGCAGCAGCTGTTTGAAGTTC 2403
|||||
Qy 2431 GCGCTGCTATCAGCAGAGTCTCTCCCTGAATGAGGCCACAGTGGCGCTCTTCGAGGTC 2490
|||||
Db 2404 TTCTGTACTTCTCTGCGTCACAGTGTCTGTAAGATTTGGTGTGACAGTGTCTATACAGTGC 2463
|||||
Qy 2491 TTCTGTACTTCTATGTGCCACACCGCTGCGCATCAGGGCAAGTCTAGTCCAGTGC 2550
|||||
Db 2464 CAGGGATCCCGGAGGCTCCAGGCTATCCACCGCTGCTCTGAGTGTGTTCGAGAGAC 2523
|||||
Qy 2551 CAGGGATCCCGGAGGCTCCATCCTCTCCACGCTGCTCTGAGGCTGTGCTACGCGGAC 2610
|||||
Db 2524 ATGAGACACAGCTGTTTGTGTCAGGTGCAGCGGATGGGTTGCTTTTACGTTTTTGTGAT 2583
|||||
```

```
RESULT 6
LOCUS AF029235 201 bp mRNA ROD 12-NOV-1997
DEFINITION Mus musculus telomerase catalytic subunit mRNA, partial cds.
ACCESSION AF029235
NID 92605902
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 201)
AUTHORS Drissi,R. and Cleveland,J.L.
TITLE Partial sequence of Mus musculus telomerase catalytic subunit
homolog
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 201)
AUTHORS Drissi,R. and Cleveland,J.L.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1997) Biochemistry, St Jude Children's Research
Hospital, 332 North Lauderdale, Memphis, TN 38105, USA
FEATURES
source 1..201
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="32D.3"
<1..>201
/codon_start=1
/product="telomerase catalytic subunit"
/db_xref="PID:g2605903"
/translation="FFVTESTFOKNRLFYKSVMSKLSQSIGVRLRLRLSLSQ
EVRHHQDTWLPICRLRFIPK"
BASE COUNT 51 a 54 c 55 g 41 t
ORIGIN
Query Match 4.5%; Score 129; DB 28; Length 201;
Best Local Similarity 82.1%; Pred. No. 1.63e-71;
Matches 165; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
Db 1 TTCTTTTATGTCAGGAGACACATTCACAGAGACAGGCTCTTCTTCTACCGTAAGAT 60
Qy 1744 TTCTTTTATGTCAGGAGACACGCTTTCAAAGAACAGGCTCTTTTCTACCGAAGAT 1803
|||||
Db 61 GTCTGGAGCAAGCTGGCAGACATTTGGATGATCAGGCAACACCTTGAGAGTGGCGCTACGG 120
Qy 1804 GTCTGGAGCAAGTTGCAAGACATTTGGAATCAGACAGCATTTGAAGAGGTGTCAGCTGCGG 1863
|||||
Db 121 GAGCTGTCAACAAGAGGAGGTTCAGGCAATCACAGGACACCTGGCTAGCGATGCCCATCTGC 180
Qy 1864 GAGCTGTGGAAGACAGAGGTTCAGGCAAGGTTCAGGCAAGGTTCAGGCAAGGTTCAGGTC 1923
|||||
Db 181 AGACTCGGCTTCATCCCAAG 201
Qy 1924 AGACTCGGCTTCATCCCAAG 1944
|||||
RESULT 7
LOCUS 166494 7218 bp DNA PAT 23-DEC-1997
```

| | |
|-----------------------|---|
| DEFINITION | Sequence 14 from patent US 5670367. |
| ACCESSION | I66494 |
| NID | G2724471 |
| KEYWORDS | . |
| SOURCE | Unknown. |
| ORGANISM | Unknown. |
| REFERENCE | Unclassified. 1 (bases 1 to 7218) |
| AUTHORS | Dorner,F., Scheiflinger,F. and Falkner,F.Gunter. |
| TITLE | Recombinant fowlpox virus |
| JOURNAL | Patent: US 5670367-A 14 23-SEP-1997; |
| FEATURES | Location/Qualifiers source 1..7218 /organism="unknown" |
| BASE COUNT | 1944 a 1491 c 1486 g 1929 t 368 others |
| ORIGIN | |
| Query Match | 2.4%; Score 67; DB 21; Length 7218; |
| Best Local Similarity | 2.4%; Pred. No. 1.13e-253; |
| Matches | 9; Conservative 211; Mismatches 15; Indels 0; Gaps 0; |
| Db | 1055 GGAGCTTGCATYY 1114 ::: : : : : ::::: ::: : : : : : : : : : : : : : |
| Qy | 921 GGAGGTGCGTCTCTGGCAGCGCCACTCCCACCATCGTGGCGGCCAGCACGCG 980 |
| Db | 1115 YY 1174 ::: |
| Qy | 981 GGGCCCCCATCCACATCGCGGCCACCACGTCCTGGGCACACGCTTGTCGCCCGGTGA 1040 |
| Db | 1175 YY 1234 ::: |
| Qy | 1041 CGCCGAGCAAGCACTTCCTCTACTCTCAGCGACAAGGACGCTGCGGCCCTCCTT 1100 |
| Db | 1235 YY 1294 : |
| Qy | 1101 CCTACTAGCTCTCTGAGGCCAGCTGACTGGCGCTCGGAGGCTCGTGAGACATCTT 1160 |
| Db | 1295 YY 1354 ::: |
| Qy | 1161 TCTGGTTCCAGGCGCTGGATGCCAGGACTCCCGCAGGTTGCCCGCCTGCCCGAGCG 1220 |
| Db | 1355 YY 1414 ::: |
| Qy | 1221 CTACTGGCAATGGCGCCCTGTTTCTTGAGACTGCTTGGGAACACGCGAGTGCCTTA 1280 |
| -Db | 1415 YYYYYYYYYYYY 1427 |
| Qy | 1281 CGGGTGCTCCTC 1293 |
| RESULT | 8 |
| LOCUS | I66494 7218 bp DNA PAT 23-DEC-1997 |
| DEFINITION | Sequence 14 from patent US 5670367. |
| ACCESSION | I66494 |
| NID | G2724471 |
| KEYWORDS | . |
| SOURCE | Unknown. |
| ORGANISM | Unknown. |
| REFERENCE | Unclassified. 1 (bases 1 to 7218) |
| AUTHORS | Dorner,F., Scheiflinger,F. and Falkner,F.Gunter. |
| TITLE | Recombinant fowlpox virus |
| JOURNAL | Patent: US 5670367-A 14 23-SEP-1997; |
| FEATURES | Location/Qualifiers source 1..7218 /organism="unknown" |
| BASE COUNT | 1944 a 1491 c 1486 g 1929 t 368 others |
| ORIGIN | |
| Query Match | 1.7%; Score 49; DB 21; Length 7218; |
| Best Local Similarity | 2.0%; Pred. No. 1.72e-13; |
| Matches | 6; Conservative 165; Mismatches 122; Indels 0; Gaps 0; |

| | |
|---------------|--|
| repeat_region | /rpt_family="Alu" 3707. .3728 /note=" (A)22" /rpt_type=tandem /rpt_unit=A 4366. .4661 /rpt_family="Alu" 5327. .5602 /rpt_family="Alu" 6586. .6956 /rpt_family="L1" 6647. .6684 /note=" (CA)19" /rpt_type=tandem /rpt_unit=CA 7113. .7373 /rpt_family="Alu" complement(7830. .8185) /standard_name="possible repeat" 8258. .8503 /rpt_family="Alu" 9070. .9387 /rpt_family="Alu" complement(9740. .9845) /rpt_family="MER42" complement(10440. .11015) /rpt_family="Alu" 11950. .12250 /rpt_family="Alu" 12037. .12085 /note=" (A)29" /rpt_type=tandem /rpt_unit=A 12365. .12645 /rpt_family="Alu" 13727. .13750 /note=" (AC)12" /rpt_type=tandem /rpt_unit=AC 13783. .14024 /rpt_family="L1" 14175. .14470 /rpt_family="Alu" complement(14906. .15259) /standard_name="possible repeat" 15300. .15613 /rpt_family="Alu" 16671. .16690 /note=" (T)20" /rpt_type=tandem /rpt_unit=T complement(16675. .16977) /rpt_family="Alu" complement(16993. .17085) /rpt_family="MER42" complement(17678. .18276) /rpt_family="Alu" 19305. .19583 /rpt_family="Alu" 19914. .19945 /note=" (GTTT)8" /rpt_type=tandem /rpt_unit=GTTT complement(19943. .20222) /rpt_family="Alu" 21202. .21496 /rpt_family="Alu" 21736. .22035 /rpt_family="Alu" 22017. .22038 /note=" (A)22" /rpt_type=tandem /rpt_unit=A |
|---------------|--|

| | |
|---------------|---|
| repeat_region | 22321. .22457 /rpt_family="Alu" complement(22638. .22981) /standard_name="possible repeat" 23022. .23326 /rpt_family="Alu" 23473. .23761 /rpt_family="Alu" 23744. .23767 /note="(A)24" /rpt_type=tandem /rpt_unit=A complement(24639. .24694) /rpt_family="MER42" complement(25349. .25713) /rpt_family="Alu" complement(25727. .26471) /rpt_family="Alu" 27191. .27477 /rpt_family="Alu" 27774. .28057 /rpt_family="Alu" 28040. .28066 /note="(A)27" /rpt_type=tandem /rpt_unit=A 28712. .28930 /rpt_family="MER20" complement(28769. .28838) /note="GRAIL 2 excellent exon, frame 0" 28987. .29214 /rpt_family="Alu" join(29485. .29559,29598. .29779) /note="95% identity dbSTS:GI4522 (SHGC-11312)" 29495. .29976 /standard_name="A1027942" /note="100% identity EST ov84a10.x1" complement(30401. .30536) /note="GRAIL 2 excellent exon, frame 2" /rpt_family="MIR" complement(30682. .30733) /rpt_family="MIR" complement(31573. .31724) /note="GRAIL 2 excellent exon, frame 1" complement(32159. .32232) /note="GRAIL 2 excellent exon, frame 2" 32386. .32488 /rpt_family="MLT1" 32617. .32908 /rpt_family="Alu" 32977. .33088 /rpt_family="MLT1" complement(33670. .33785) /rpt_family="Alu" complement(34021. .34144) /note="GRAIL 2 excellent exon, frame 2" complement(35238. .35331) /note="GRAIL 2 excellent exon, frame 1" 36392. .36663 /rpt_family="Alu" 36901. .37222 /note="100% identity EST ou55c09.x1" /db_xref="dbEST:A1025011" 36901. .37164 /standard_name="STRSG-9983" /db_xref="dbSTS:G36554" complement(37595. .37654) /note="GRAIL 2 excellent exon, frame 0" complement(join(38063. .38218,38462. .38578,3839071. .39205,39532. .39630,39935. .40048,403040503. .40661,41868. .41972,42103. .42225,424944379. .44507)) /standard_name="histidyl-trNA synthetase" /note="78%-100% protein identity GenPept:U1R93" |
|---------------|---|

•


```
misc_feature complement(38069..38215)
/note="GAIL 2 excellent exon, frame 0"
misc_feature complement(38462..38578)
```

```
misc_feature
complement(38462..38578)
/note= GRAB 2 excellent
```

```
repeat_region /standard_name="RLF"
repeat_region /note="65% & 69% protein identity GenPept:U22377"
repeat_region 3431..3724
repeat_region /rpt_family="Alu"
repeat_region 3707..3728
repeat_region /note="(A)22"
repeat_region /rpt_type=tandem
repeat_region /rpt_unit=A
repeat_region 4366..4661
repeat_region /rpt_family="Alu"
repeat_region 5327..5602
repeat_region /rpt_family="Alu"
repeat_region 6586..6956
repeat_region /rpt_family="L1"
repeat_region 6647..6684
repeat_region /note="(CA)19"
repeat_region /rpt_type=tandem
repeat_region /rpt_unit=CA
repeat_region 7113..7373
repeat_region /rpt_family="Alu"
repeat_region /standard_name="possible repeat"
repeat_region 8258..8503
repeat_region /rpt_family="Alu"
repeat_region 9070..9387
repeat_region /rpt_family="Alu"
repeat_region /complement(9740..9845)
repeat_region /rpt_family="MER42"
repeat_region /complement(10440..11015)
repeat_region 11950..12250
repeat_region /rpt_family="Alu"
repeat_region 12057..12085
repeat_region /note="(A)29"
repeat_region /rpt_type=tandem
repeat_region /rpt_unit=A
repeat_region 12365..12645
repeat_region /rpt_family="Alu"
repeat_region 13727..13750
repeat_region /note="(AC)12"
repeat_region /rpt_type=tandem
repeat_region /rpt_unit=AC
repeat_region 13783..14024
repeat_region /rpt_family="L1"
repeat_region 14175..14470
repeat_region /rpt_family="Alu"
repeat_region /complement(14906..15259)
repeat_region /standard_name="possible repeat"
repeat_region 15300..15613
repeat_region /rpt_family="Alu"
repeat_region 16671..16690
repeat_region /note="(T)20"
repeat_region /rpt_type=tandem
repeat_region /rpt_unit=T
repeat_region /complement(16675..16977)
repeat_region /rpt_family="Alu"
repeat_region /complement(16993..17085)
repeat_region /rpt_family="MER42"
repeat_region /complement(17678..18276)
repeat_region /rpt_family="Alu"
repeat_region 19305..19583
repeat_region /rpt_family="Alu"
repeat_region 19914..19945
repeat_region /note="(GTTT)8"
repeat_region /rpt_type=tandem
repeat_region /rpt_unit=TTTT
repeat_region /complement(19943..20222)
repeat_region /rpt_family="Alu"
repeat_region 21202..21496
repeat_region /rpt_family="Alu"
repeat_region 21736..22035
repeat_region /rpt_family="Alu"
repeat_region 22017..22038
```

```
repeat_region /note="(A)22"
repeat_region /rpt_type=tandem
repeat_region 22321..22457
repeat_region /rpt_family="Alu"
repeat_region /complement(22638..22981)
repeat_region /standard_name="possible repeat"
repeat_region 23022..23326
repeat_region /rpt_family="Alu"
repeat_region 23473..23761
repeat_region /rpt_family="Alu"
repeat_region 23744..23767
repeat_region /note="(A)24"
repeat_region /rpt_type=tandem
repeat_region /rpt_unit=A
repeat_region /complement(24639..24694)
repeat_region /rpt_family="MER42"
repeat_region /complement(25349..25713)
repeat_region /rpt_family="Alu"
repeat_region /complement(25727..26471)
repeat_region 27191..27477
repeat_region /rpt_family="Alu"
repeat_region 27774..28057
repeat_region /rpt_family="Alu"
repeat_region 28040..28066
repeat_region /note="(A)27"
repeat_region /rpt_type=tandem
repeat_region /rpt_unit=A
repeat_region 28712..28930
repeat_region /rpt_family="MER20"
repeat_region /complement(28769..28838)
repeat_region /note="GRAIL 2 excellent exon, frame 0"
repeat_region 28987..29214
repeat_region /rpt_family="Alu"
repeat_region /note="95% identity dbSTS:G14522 (SHGC-11312)"
repeat_region 29495..29976
repeat_region /standard_name="AI027942"
repeat_region /note="100% identity EST ov84a10.x1"
repeat_region /complement(30401..30536)
repeat_region /note="GRAIL 2 excellent exon, frame 2"
repeat_region /complement(30682..30733)
repeat_region /rpt_family="Mir"
repeat_region /complement(31573..31724)
repeat_region /note="GRAIL 2 excellent exon, frame 1"
repeat_region /complement(32159..32232)
repeat_region /note="GRAIL 2 excellent exon, frame 2"
repeat_region 32388..32488
repeat_region /rpt_family="MLT1"
repeat_region 32617..32908
repeat_region /rpt_family="Alu"
repeat_region 32977..33088
repeat_region /rpt_family="MLT1"
repeat_region /complement(33670..33785)
repeat_region /rpt_family="Alu"
repeat_region /complement(34021..34144)
repeat_region /note="GRAIL 2 excellent exon, frame 2"
repeat_region /complement(35238..35331)
repeat_region /note="GRAIL 2 excellent exon, frame 1"
repeat_region 36392..36663
repeat_region /rpt_family="Alu"
repeat_region 36901..37222
repeat_region /note="100% identity EST ou55c09.x1"
repeat_region /db_xref="dbEST:AI025011"
repeat_region 36901..37164
repeat_region /standard_name="STSG-9983"
repeat_region /db_xref="dbSTS:G26554"
repeat_region /complement(37595..37654)
repeat_region /note="GRAIL 2 excellent exon, frame 0"
repeat_region /complement(join(38063..38218,38462..38578,38741..38995,
39071..39205,39532..39630,39935..40048,40300..40410,
40503..40661,41868..41972,42103..42225,42492..42569,
```


Cp 49 CCAGGGCTTCCACGTCGCGAGCA 26

Search completed: Fri Dec 25 12:03:06 1998
Job time : 11319 secs.

W P E R E L F (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 12:03:25 1998; MasPar time 8218.69 Seconds
Tabular output not generated. 620.553 Million cell updates/sec

Title: >US-08-951-733-13
Description: (1-2848) from US08951733.seq (1 of 3)
Perfect Score: 2848
N.A. Sequence: 1 CACGCGTCGGCGAGCGCTG.....GATCGCGGCGCCAGCGCTAT 2848
Comp: GTGCGGAGCGCGCTGCGGAC.....CTACGCGCGGTCGCGGATA

Scoring table: TABLE default
Gap open 30; Gap extend 1
Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl-est55

Database: genbank-est107
1:em_est1 2:em_gss1 3:em_gss2 4:em_gss3
5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est20 17:gb_est21
18:gb_est21 19:gb_est23 20:gb_est4 21:gb_est5 22:gb_est6
23:gb_est7 24:gb_est8 25:gb_est9 26:gb_gss1 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 11.733; Variance 2.670; scale 4.395

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|-------------|------------------------|
| 1 | 385 | 13.5 | 389 | 8 | AA281296 | zt08g02.rl NCI_CGAP_GC |
| 2 | 187 | 6.6 | 409 | 24 | AA311750 | EST182469 Jurkat T-cel |
| 3 | 58 | 2.0 | 252 | 12 | AA754459 | 97SN1787 Rice Immature |
| 4 | 54 | 1.9 | 252 | 12 | AA754459 | 97SN1787 Rice Immature |
| 5 | 49 | 1.7 | 247 | 12 | AA754458 | 97SN1784 Rice Immature |
| 6 | 44 | 1.5 | 247 | 12 | AA754458 | 97SN1784 Rice Immature |
| 7 | 33 | 1.2 | 375 | 23 | AA200728 | mul3h09.rl Soares 2BDM |
| 8 | 33 | 1.2 | 2275 | 11 | AF034173 | Homo sapiens ntcon2 co |
| 9 | 31 | 1.1 | 660 | 11 | AF034177 | Homo sapiens ntcon6 co |
| 10 | 28 | 1.0 | 1287 | 12 | AF038250 | Homo sapiens clone ntc |
| 11 | 28 | 1.0 | 1287 | 12 | AF038250 | Homo sapiens clone ntc |
| 12 | 25 | 0.9 | 143 | 18 | AA117394 | EST217350 Normalized r |

| | | | | | | | |
|----|----|-----|------|----|-----------|------------------------|----------|
| 13 | 25 | 0.9 | 181 | 6 | AA386387 | EST81369 Prostate glan | 8.99e-04 |
| 14 | 25 | 0.9 | 187 | 13 | AA855630 | Y06h11.rl Stratagene | 8.99e-04 |
| 15 | 25 | 0.9 | 213 | 16 | R72797 | Y109c09.rl Homo sapien | 8.99e-04 |
| 16 | 25 | 0.9 | 242 | 15 | AA1012107 | EST206558 Normalized r | 8.99e-04 |
| 17 | 25 | 0.9 | 278 | 13 | AA882818 | TENS0393 T. cruzi epim | 8.99e-04 |
| 18 | 25 | 0.9 | 301 | 14 | AA848961 | EST191723 Normalized r | 8.99e-04 |
| 19 | 25 | 0.9 | 302 | 15 | AA1010844 | EST205295 Normalized r | 8.99e-04 |
| 20 | 27 | 0.9 | 317 | 27 | AQ080007 | CIT-HSP-2288C3.Tf CIT- | 6.33e-06 |
| 21 | 25 | 0.9 | 318 | 14 | AA799774 | EST189271 Normalized r | 8.99e-04 |
| 22 | 25 | 0.9 | 318 | 16 | H50134 | Y027a07.rl Homo sapien | 8.99e-04 |
| 23 | 25 | 0.9 | 336 | 18 | AA1172396 | EST218402 Normalized r | 8.99e-04 |
| 24 | 26 | 0.9 | 343 | 22 | AA098755 | T3985 MVAT4 bloodstrea | 7.82e-05 |
| 25 | 25 | 0.9 | 359 | 24 | AA303595 | EST16305 Aorta endothe | 8.99e-04 |
| 26 | 25 | 0.9 | 382 | 5 | RS4856 | YJ74d04.rl Homo sapien | 8.99e-04 |
| 27 | 26 | 0.9 | 396 | 19 | T26788 | T529 Trypanosoma bruce | 7.82e-05 |
| 28 | 26 | 0.9 | 400 | 16 | H44707 | YP24c08.rl Homo sapien | 7.82e-05 |
| 29 | 25 | 0.9 | 420 | 17 | AI071447 | UI-R-C1-ku-c-05-0-UI.s | 8.99e-04 |
| 30 | 25 | 0.9 | 438 | 14 | AA924206 | UI-R-Al-dx-a-11-0-UI.s | 8.99e-04 |
| 31 | 25 | 0.9 | 478 | 14 | AA925372 | UI-R-Al-ee-d-07-0-UI.s | 8.99e-04 |
| 32 | 25 | 0.9 | 487 | 16 | R86860 | Ym86a03.rl Homo sapien | 8.99e-04 |
| 33 | 25 | 0.9 | 504 | 15 | AI010120 | EST204571 Normalized r | 8.99e-04 |
| 34 | 25 | 0.9 | 511 | 17 | AI072026 | UI-R-C2-nd-e-12-0-UI.s | 8.99e-04 |
| 35 | 25 | 0.9 | 511 | 15 | AA997735 | UI-R-C0-hn-b-08-0-UI.s | 8.99e-04 |
| 36 | 25 | 0.9 | 540 | 14 | AA901048 | UI-R-Al-dv-e-09-0-UI.s | 8.99e-04 |
| 37 | 25 | 0.9 | 590 | 10 | AA622227 | no43c05.sl NCI_CGAP.Pr | 8.99e-04 |
| 38 | 27 | 0.9 | 660 | 11 | AF034177 | Homo sapiens ntcon6 co | 6.33e-06 |
| 39 | 25 | 0.9 | 677 | 18 | AI105428 | EST214717 Normalized r | 8.99e-04 |
| 40 | 25 | 0.9 | 705 | 14 | AA941159 | LD25143.5prime LD Dros | 8.99e-04 |
| 41 | 26 | 0.9 | 7275 | 11 | AF034173 | Homo sapiens ntcon2 co | 7.82e-05 |
| 42 | 24 | 0.8 | 299 | 21 | W04101 | TL794 MVAT4 bloodstrea | 9.56e-03 |
| 43 | 24 | 0.8 | 398 | 5 | T47934 | Yb18d04.sl Homo sapien | 9.56e-03 |
| 44 | 24 | 0.8 | 456 | 6 | CI9919 | Rice cDNA, partial seq | 9.56e-03 |
| 45 | 24 | 0.8 | 901 | 20 | W21990 | 60h10 Human retina cDN | 9.56e-03 |

ALIGNMENTS

| RESULT | 1 | AA281296 | 389 bp | mrna | EST | 14-AUG-1997 |
|------------|---|--|--------------|-------------------------|-----|-------------|
| LOCUS | | zt08g02.rl NCI_CGAP_GCBI | Homo sapiens | cDNA clone IMAGE:712562 | 5', | |
| DEFINITION | | mrna sequence. | | | | |
| ACCESSION | | AA281296 | | | | |
| NID | | G1924194 | | | | |
| KEYWORDS | | EST. | | | | |
| SOURCE | | human. | | | | |
| ORGANISM | | Homo sapiens | | | | |
| REFERENCE | | 1 (bases 1 to 389) | | | | |
| AUTHORS | | NCI-CGAP | | | | |
| TITLE | | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), | | | | |
| JOURNAL | | Tumor Gene Index | | | | |
| COMMENT | | Unpublished (1997) | | | | |

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 2187 Std Error: 0.00

Seq primer: -28m13 rev2 Et from Amersham

High quality sequence stop: 385.

Location/Qualifiers

source

/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD+),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer

[5'-TGTTACCAATCTCAAGTGGAGCGCGCTCATTTTTTTTTTTTTTT-3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

/db_xref="taxon:9606"
/clone="IMAGE:712562"
/clone_lib="NCI_GCAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
<1. ->389

BASE COUNT 87 a 102 c 123 g 77 t
ORIGIN

Query Match 13.5%; Score 385; DB 8; Length 389;

Best Local Similarity 99.5%; Pred. No. 0.00e+00;

Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 GCCAAGTTCCTGCACCTGCTGATGAGTGTACGTGCTGAGCTGCTCAGTCTTTCTTT 60
QY 1690 GCCAAGTTCCTGCACCTGCTGATGAGTGTACGTGCTGAGCTGCTCAGTCTTTCTTT 1749

Db 61 TATGTCAGGAGACACACCTTTCAAAGAACAGGCTCTTTTCTACCGGAAGTGTCTGG 120

QY 1750 TATGTCAGGAGACACACCTTTCAAAGAACAGGCTCTTTTCTACCGGAAGTGTCTGG 1809

Db 121 AGCAAGTTGCAAGCATTGGAATCAGACAGCACCTTGAAGAGGGTGCAGCTCGGGAGCTG 180

QY 1810 AGCAAGTTGCAAGCATTGGAATCAGACAGCACCTTGAAGAGGGTGCAGCTCGGGAGCTG 1869

Db 181 TCGGAGCAGAGGTGAGCAGCATCGGGAAGCCAGGCGCCGCTGCTGACGTCAGACTC 240

QY 1870 TCGGAGCAGAGGTGAGCAGCATCGGGAAGCCAGGCGCCGCTGCTGACGTCAGACTC 1929

Db 241 CGCTTCATCCCAAGCCTGACGGCTCGCGCGGATTGTGAATGGACTAGCTGCTGGGA 300

QY 1930 CGCTTCATCCCAAGCCTGACGGCTCGCGCGGATTGTGAATGGACTAGCTGCTGGGA 1989

Db 301 GCCAAGCTTCCCGCAGAAAAGAGGCCGAGCGCTCTCACCTCGAGGGTGAAGGCACGTG 360

QY 1990 GCCAAGCTTCCCGCAGAAAAGAGGCCGAGCGCTCTCACCTCGAGGGTGAAGGCACGTG 2049

Db 361 TTCACGTGCTCACTACGACGGCGCGG 389

QY 2050 TTCACGTGCTCACTACGACGGCGCGG 2078

RESULT 2 AA311750 409 bp mRNA EST 19-APR-1997
LOCUS EST182469 Jurkat T-cells Vi Homo sapiens cDNA 5' end, mRNA
DEFINITION sequence.

ACCESSION AA311750

NID g1964077

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae;

Homo.

1 (bases 1 to 409)

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,

He,W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kusch,C., Hungjun,J., Li,H., Meissner,P., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence Nature 377 (6547 Suppl), 3-174 (1995)

JOURNAL MEDLINE COMMENT

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/ngi/ngi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

source

1. .409

/organism="Homo sapiens"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:

XhoI"

/db_xref="ATCC (inhost):158964"

/db_xref="taxon:9606"

/clone_lib="Jurkat T-cells VI"

/cell_type="T-lymphocyte"

<1. ->409

BASE COUNT 65 a 120 c 133 g 86 t 5 others

ORIGIN

Query Match 6.6%; Score 187; DB 24; Length 409;

Best Local Similarity 98.9%; Pred. No. 5.02e-269;

Matches 187; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 GTTTGGTGGATGATTTCTTTGGTGACACTCACCTCACCCACGCGAAAACCTTCCTCA 60

QY 2660 GTTTGGTGGATGATTTCTTTGGTGACACTCACCTCACCCACGCGAAAACCTTCCTCA 2719

Db 61 GGACCTTGTCGGAGGTGCTCCCTGAGTATGGCTCGGTGGTGAACCTTGGGAGACAGTGG 120

QY 2720 GGACCTTGTCGGAGGTGCTCCCTGAGTATGGCTCGGTGGTGAACCTTGGGAGACAGTGG 2779

Db 121 TGAACCTCCCTGTAGAAGACGAGCCCTGGGTGCACGGCTTTTTCAGATGCGCGCC 180

QY 2780 TGAACCTCCCTGTAGAAGACGAGCCCTGGGTGCACGGCTTTTTCAGATGCGCGCC 2839

Db 181 ACGCCTAT 189

QY 2840 ACGCCTAT 2848

RESULT 3

LOCUS

DEFINITION

AA754459 252 bp mRNA EST

97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa

ACCESSION

AA754459

NID

g2801165

KEYWORDS

EST.

SOURCE

ORGANISM

Oryza sativa

rice.

REFERENCE

1 (bases 1 to 252)

AUTHORS

Nahm,B.H., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,

Lee,M.C. and Eun,M.Y.

Large-scale Sequencing Analysis of ESTs from Rice Immature Seed

JOURNAL

Unpublished (1998)

Eukaryotae; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.


```
Query Match 1.2%; Score 33; DB 23; Length 375;
Best Local Similarity 75.4%; Pred. No. 5.05e-13;
Matches 49; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Db 144 AGAGCATCTATGAATGAGAGCAGCAGCCTGTTTGACTTCTTCTGCACTTCCTGC 203
||||| ||| ||||||| ||||||| ||||| ||| ||| ||||||| |||||||
Qy 2447 AGAGCTCTCCTGAATGAGGCCAGCAGTGGCTTTCGAGCTCTTCTACGCTCAGT 2506
Db 204 GTCAC 208
Qy 2507 GCCAC 2511

RESULT 8 AF034173 2275 bp mRNA EST 22-DEC-1997
LOCUS Homo sapiens ntcon2 contig mRNA, partial sequence, mRNA sequence.
ACCESSION AF034173
NID 92707735
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2275)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Generation of a transcription map in the region immediately
centromeric to human MHC across the 6p21.2-6p21.3 chromosomal
boundary
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2275)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1997) Division of Medical and Molecular Genetics,
Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK
FEATURES
source
1. 2275
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/clone="ntcon2 contig"
/tissue_type="liver; brain"
/dev_stage="fetus"
/note="similar to Br140"

BASE COUNT 438 a 619 c 470 g 599 t 149 others
ORIGIN

Query Match 1.2%; Score 33; DB 11; Length 2275;
Best Local Similarity 13.3%; Pred. No. 5.05e-13;
Matches 11; Conservative 47; Mismatches 25; Indels 0; Gaps 0;

Db 1523 RGMKCKYKRYKYKYSTYKSWRYWYTYTYWCCTSMKASGASCMRMWGMGYSR 1582
:||||: :|||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 2648 GGCTGCTCTGCTGGTTGGTGATGATTCTTGTGGTGCACCTCACCACCGGGA 2707
Db 1583 SRSRWYWGWSMGCGYMTKRY 1605
Qy 2708 AAACCTTCTCAGGACCCCTGGTC 2730

RESULT 9 AF034177 660 bp mRNA EST 22-DEC-1997
LOCUS Homo sapiens ntcon6 contig mRNA, partial sequence, mRNA sequence.
ACCESSION AF034177
NID 92707739
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Generation of a transcription map in the region immediately
```

```
centromeric to human MHC across the 6p21.2-6p21.3 chromosomal
boundary
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 660)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1997) Division of Medical and Molecular Genetics,
Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK
FEATURES
source
1. 660
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/clone="ntcon6 contig"
/tissue_type="fetal brain; fetal liver; adult muscle"
/note="similar to CutA"

BASE COUNT 162 a 161 c 173 g 109 t 55 others
ORIGIN

Query Match 1.1%; Score 31; DB 11; Length 660;
Best Local Similarity 15.6%; Pred. No. 1.45e-10;
Matches 7; Conservative 31; Mismatches 7; Indels 0; Gaps 0;

Db 384 WSSRCYRSYGYHYCYKCDMSCTTSKSRWYKSRSWCDD 428
: : ||| : : : : : : : : : : : : : : : : : : : : : : : :
Qy 347 AGAAGTGCTGGCTTCGGCTTCGGCTGCTGGAGGGGCGCG 391

RESULT 10 AF038250 1287 bp mRNA EST 29-JAN-1998
LOCUS Homo sapiens clone ntcon9 mRNA, partial sequence, mRNA sequence.
ACCESSION AF038250
NID 92815880
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1287)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-1997) Division of Medical and Molecular Genetics,
Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK
FEATURES
source
1. 1287
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/clone="ntcon9"
/dev_stage="fetal"
/tissue_type="brain; liver"
/note="similar to HSRP20"

BASE COUNT 349 a 219 c 293 g 361 t 65 others
ORIGIN

Query Match 1.0%; Score 28; DB 12; Length 1287;
Best Local Similarity 29.1%; Pred. No. 4.79e-07;
Matches 23; Conservative 29; Mismatches 27; Indels 0; Gaps 0;

Db 401 DCBTGGDPTBSHBVCMBCBANADGATBACGKGVGBGASTCMTHYNCCDCKTGSAG 460
:||||: :|||: : : : : : : : : : : : : : : : : : : : : : : :
Qy 227 GCTGTGTGCTGCTGGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 286
Db 461 TVNHDWSMAGARAGCT 479
Qy 287 TGCTCTGCTGAAGGAGCT 305

RESULT 11 AF038250 1287 bp mRNA EST 29-JAN-1998
LOCUS Homo sapiens clone ntcon9 mRNA, partial sequence, mRNA sequence.
```

ORIGIN

Query Match 0.9%; Score 25; DB 18; Length 143;
 Best Local Similarity 76.6%; Pred. No. 8.99e-04;
 Matches 36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Db 4 GGGAGCGCATGGCCGCTGCGGGGCTCGGCGCCGCTCCGCTGGTG 50
 ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Cp 103 GGGAGCGCATGGCTGCGAGCGGGAGCGCGGCATCGCGGGGTG 57

RESULT 13

LOCUS AA386387 181 bp mRNA EST 22-APR-1997
 DEFINITION EST01369 Prostate gland i Homo sapiens cDNA 5' end similar to
 prol14 4-hydroxylase, beta subunit/protein disulfide
 isomerase/thyroid hormone-binding protein, mRNA sequence.

ACCESSION AA386387
 NID 92038747
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
 Homo.

REFERENCE 1 (bases 1 to 181)
 AUTHORS Adams,M.D., Kervage,A.R., Fleischmann,R.D., Fuldner,R.A.,
 Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.B., Geoghagen,N.S.,
 Glodok,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
 He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
 Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
 Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M.,
 Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
 Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence

JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
 MEDLINE 96026280
 COMMENT Other_ESTs: EST81368 THCL170792
 Contact: Kervage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org

For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/cdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES
 source Location/Qualifiers
 1..181
 /organism="Homo sapiens"
 /note="Organ: prostate; Vector: pBluescript SK-; Site_1:
 ECORI; Site_2: XhoI"
 /db_xref="ATCC (inhost):174439"
 /db_xref="taxon:9606"
 /clone_lib="Prostate gland I"
 /sex="male"
 /dev_stage="adult, 21 yrs"
 <1..>181 32 t 5 others

BASE COUNT 25 a 60 c 59 g 32 t
 ORIGIN

Query Match 0.9%; Score 25; DB 6; Length 181;

```
Best Local Similarity 70.0%; Pred. No. 8.99e-04;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Db 32 TGCCCGCGCTCTCTGTGCTGCGTGGCGCCGCGCGCGCGCGAGACCCNAGG 91
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 68 TGCCCGCGCTCCCGCTGCGGAGCGGTGCGCTCCCTGCTGCGGAGCAGCTACCGCGAGG 127

RESULT 14
LOCUS AA855630 187 bp mRNA EST 06-MAR-1998
DEFINITION vv68h11.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone
1260165 5' similar to TR:Q99960 Q99960 PLAKOPHILIN 2A. [1] ;, mRNA
sequence.
ACCESSION AA855630
NID 92943168
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 187)
AUTHORS Marra,M., Hallier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:662717
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 178.
FEATURES
source 1..187
Location/Qualifiers
/organism="Mus musculus"
/strain="NIH/Swiss"
/notes="Organ: heart; Vector: pBluescript SK-; Site.1:
EcoRI; Site.2: XhoI; Cloned unidirectionally. Primer:
Oligo dT: 93 pooled NIH/Swiss 13 day embryo hearts.
Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor
sequence: 5' CTCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3'"
/db_xref="taxon:10090"
/clone="1260165"
/clone_lib="Stratagene mouse heart (#937316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
BASE COUNT 29 a 65 c 70 g 23 t
ORIGIN
Query Match 0.9%; Score 25; DB 13; Length 187;
Best Local Similarity 85.7%; Pred. No. 8.99e-04;
Matches 30; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 105 GCCTGGCGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 169 GCCTGGCGGCTGGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203

RESULT 15
LOCUS R72797 213 bp mRNA EST 02-JUN-1995
DEFINITION y109c09.r1 Homo sapiens cDNA clone 157744 5'.
```

```
ACCESSION R72797
NID 9846829
KEYWORDS EST.
SOURCE human clones-157744 library=Soares breast 2NbHst vector=pt7T3D
(Pharmacia) with a modified polylinker host=DH10B (ampicillin
resistant) primer=M13p1 Rsite1=Not I Rsite2=Eco RI Adult female.
1st strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCGAATGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia),
digested with Not I and cloned into the Not I and Eco RI sites of a
modified p7T3 vector (Pharmacia). Library went through one round
of normalization to a Cot = 230. Library constructed by Bento
Soares and M.Fatima Bonaldo.
Homo sapiens
ORGANISM Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 213)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasaki,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
High quality sequence stops: 185
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES
source 1..213
Location/Qualifiers
/organism="Homo sapiens"
/clone="157744"
BASE COUNT 40 a 72 c 69 g 32 t
ORIGIN
Query Match 0.9%; Score 25; DB 16; Length 213;
Best Local Similarity 74.5%; Pred. No. 8.99e-04;
Matches 38; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Db 32 CTGCCCCGGCGCCAGCTTTGTGAGCAGCGTGAGACCCGAGAGCTCTCGGCTG 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 CTGCGGTGGCGCAGCTGTCGTGCGCGCTGCGGCGCCCGCGCGCTGCGGCTG 180

Search completed: Fri Dec 25 14:50:06 1998
Job time : 10001 secs.
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

(TM)

| Result | No. | Score | Query | | Length | DB | ID | Description | Pred. No. |
|--------|-----|-------|-------|-----|--------|--------|-----------------------|-------------|-----------|
| | | | Match | % | | | | | |
| C | 1 | 45 | 1.6 | 91 | 9 | Q51746 | Oligonucleotide probe | 1.95e-10 | |
| | 2 | 45 | 1.6 | 91 | 9 | Q51746 | Oligonucleotide probe | 1.95e-10 | |
| | 3 | 43 | 1.5 | 204 | 1 | N81164 | Base substituted E.co | 2.32e-09 | |
| | 4 | 40 | 1.4 | 114 | 12 | Q70466 | Generic DNA sequence | 8.94e-08 | |
| | 5 | 39 | 1.4 | 114 | 12 | Q70469 | Generic DNA sequence | 2.97e-07 | |
| C | 6 | 39 | 1.4 | 178 | 32 | T76405 | Human endothelin-1 an | 2.97e-07 | |
| | 7 | 41 | 1.4 | 204 | 1 | N81164 | Base substituted E.co | 2.67e-08 | |
| | 8 | 38 | 1.3 | 114 | 12 | Q70468 | Generic DNA sequence | 9.79e-07 | |
| | 9 | 38 | 1.3 | 114 | 12 | Q70465 | Generic DNA sequence | 9.79e-07 | |
| | 10 | 36 | 1.3 | 114 | 12 | Q70467 | Generic DNA sequence | 1.03e-05 | |
| | 11 | 36 | 1.3 | 114 | 12 | Q70470 | Generic DNA sequence | 1.03e-05 | |
| | 12 | 38 | 1.3 | 114 | 12 | Q70465 | Generic DNA sequence | 9.79e-07 | |
| | 13 | 38 | 1.3 | 114 | 12 | Q70465 | Generic DNA sequence | 9.79e-07 | |

Best Local Similarity 20.9%; Pred. No. 2.32e-09;
Matches 31; Conservative 57; Mismatches 60; Indels 0; Gaps 0;

Db 45 bcyrraggnycccccgggwyccgagcgcgaayydcchvgcgymrttthhyrmbnbvyrdd 104
: : : : : ||| | : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 829 TCTGTGGCCACCGGCGAGGAGCGGTGCAGCAGTACCCTGTTCGTGTGTGTGTCA 888
: : : : : ||| | : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 105 ynrsdaawycyyrryskvydcynachdhdybbvynvhnnncnccbnnhvchnv 164
: : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| :
QY 889 CTTGCCAGACCGCGGAGAGACCACTTCTTGAGGGTGGCTCTCTGGCACGCGCCAC 948
: : : : : ||| | : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 165 hbnhnwnayvrhdarddvhocvhcc 192
: : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| :
QY 949 TCCCACCATCGTGGCGCGCAGCAC 976

RESULT 4 Location/Qualifiers
ID Q70466 standard; DNA; 114 BP.
AC Q70466;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR-9 peptide library.
KW TSAR: Total synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct, rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
Key Key
FH misc_feature 55..60
FT FT
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments")
FT W09418318-A.
PN 18-AUG-1994.
PD 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
DR P-PSDB; R65152.
DR Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.

CC Q70466 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.this generic formula can also be
CC represented as follows:X(NNB)1(TGC)(NNB)10(TGC)2(NNB)4Z(NNB)8(TGC)(NNB)
CC -9Y. X and Y are flanking restriction sites (X is not the same as Y)
CC that are not specified further. Other generic sequences are shown in
CC Q70466-68. Other specific peptides generated by these generic sequences
CC are shown in R65151-54. TSARs are concatenated heterofunctional proteins
CC or peptides, comprising at least two functional regions - a binding
CC domain with affinity for a ligand and a second effector peptide portion
CC that is chemically or biologically active.They may further comprise a
CC linker peptide between the 2 domains.The oligonucleotides are also
CC designed so that the expressed peptide contains 2 or 4 cysteine residues
CC positioned in, or flanking, the unpredicted or variant residues. These
CC residues confer some degree of conformational rigidity to the peptides.
CC The TSARs or compsns. comprising a TSAR binding domain can be used in
CC vivo to deliver a chemically or biologically active moiety, eg. metal
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC on the cell. They can also replace the function of macromolecules, eg. for
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.
CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 1.4%; Score 40; DB 12; Length 114;
Best Local Similarity 8.9%; Pred. No. 8.94e-08;
Matches 10; Conservative 32; Mismatches 70; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 148 | GTGGCGCGCTGGGGCCCCAGGCGTGGCGGCTGTGACGCGGGGACCCGGCGGGTTTC | 207 |
| | | | |
| Db | 63 | bnnbnnbnnbnnbnnbnnbTgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb | 114 |
| | | : : : : : | |
| QY | 208 | CGCGCGCTGTGGCCCCAGTGCCTCGTGTGTGGCTGCCCTGGGACGACGCGGGCGC | 259 |

| | |
|----|--|
| DE | Generic DNA sequence to generate a random TSAR peptide library. |
| KW | TSAR; totally synthetic affinity reagent; synthetic; binding domain; |
| KW | effector domain; concatenated heterofunctional protein; linker; |
| KW | direct; rapid; detection; screening; treatment; generic; ss. |
| OS | Synthetic. |
| FH | Key |
| FT | Location/Qualifiers |
| FT | misc_feature 55..60 |
| FT | /tag= a |
| FT | /note= "this sequence represents '2'; 2 can be a |
| FT | sequence of 6,9 or 12 nucleotides (see |

| IN | Comments) |
|--------------|------------------------------|
| FN | |
| WO9418318-A. | |
| PD | 18-AUG-1994. |
| PF | 01-FEB-1994. U000977. |
| PR | 01-FEB-1993; US-013416. |
| PR | 30-DEC-1993; US-176500. |
| PR | 31-JAN-1994; US-189331. |
| PA | (UYNC-) UNIV NORTH CAROLINA. |
| PI | Fowlkes DM, Kay BK. |
| DR | WPI: 94-279739/34. |

Identifying proteins or peptide(s) which bind a ligand - by screening a recombinant vector library expressing fusion proteins comprising a binding domain and an effector domain

Disclosure; Page 35; 255pp; English.

Q70469 is a generic DNA sequence used to generate random TSAR peptide

This generic formula can be represented as follows: X(TGC)(NNB)10-(TGC)(NNB)6Z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction sites (X is not the same as Y) that are not specified further. This sequence generates peptides that are cloverleaf in structure. Other generic sequences are shown in Q70465-68. Other specific peptides generated by these generic sequences are shown in R5150-54. TSARs are concatenated heterofunctional proteins or peptides, comprising at least two functional regions - a binding domain with affinity for a ligand and a second effector peptide portion that is chemically or biologically active. They may further comprise a linker peptide between the 2 domains. The oligonucleotides are also designed so that the expressed peptide contains 2 or 4 cysteine residues positioned in, or flanking, the conformational rigidity to the peptides. The TSARs or compsns. comprising a TSAR binding domain can be used in vivo to deliver a chemically or biologically active moiety, eg. metal ion, radioisotope, peptide, toxin or enzyme, to the specific target or on the cell. They can also replace the function of macromolecules, eg. monoclonal or polyclonal antibodies and therefore circumvent the need for complex methods of hybridoma formation or in vivo antibody production. The TSARs are easily characterised and have designed activity allowing direct and rapid detection in a screening process.

Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

| | | | | |
|-----------------------|-------|---------------------|--------|----------------|
| Query Match | 1.48; | Score 39; | DB 12; | Length 114; |
| Best Local Similarity | 7.28; | Pred. NO. 2.97e-07; | | |
| Matches | 8; | Conservative | 32; | Mismatches 71; |
| | | | Indels | 0; |
| | | | Gaps | 0; |

[illegible][illegible]

```

RESULT      7
ID          N81164 standard; DNA; 204 BP.
AC          N81164;
DT          08-NOV-1990 (first entry)
DE          Base substituted E.coli beta-galactosidase alpha-fragment.
KW          E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS          Escherichia coli.
KEY         Location/Qualifiers
FT          misc_feature      19..69
FT          /tag= a
FT          /function= multiple cloning site
FT          primer_bind      187..204
FT          /tag= b
PN          EP-285123-A.
PD          05-MAY-1988.
PF          30-MAR-1988; 105163.
PR          03-APR-1987; US-034819.
PA          (SUSO) SUOMEN SOKERI OY.
PI          Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
PR          WPI; 88-279927/40.
DR          Introducing random point mutations into nucleic acids -
PT          by prepn of single stranded template, annealing a primer, elongation,
PT          misincorporation, completion of molecules and screening.
PS          Disclosure; p; English.
CC          Random point mutations were introduced into the alpha fragment of
CC          E.coli beta-galactosidase. The wild type sequence was obtained as a

```


DE Generic DNA sequence to generate a random TSAR peptide library.
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; generic; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 55..60
 FT /tag= a
 FT /note= "this sequence represents 'Z'; 2 can be a
 FT sequence of 6, 9 or 12 nucleotides (see
 FT comments)"
 PN WO9418318-A.
 PD 18-AUG-1994.
 PE 01-FEB-1994; U00977.
 PR 01-FEB-1993; US-013416.
 PR 30-DEC-1993; US-176500.
 PR 31-JAN-1994; US-189331.
 PA (UUNC-) UNIV NORTH CAROLINA.
 PI Fowlkes DM, Kay BK;
 DR WPI: 94-279739/34.
 DR P-PSDB; R65154.
 DR Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PT comprising a binding domain and an effector domain
 PS Disclosure; Page 35; 255pp; English.
 CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be
 CC represented as follows: X(NNB)6(TGC)(NNB)12(NNB)14(TGC)(NNB)3Y. X
 CC and Y are flanking restriction sites (X is not the same as Y) that are
 CC not specified further. Other generic sequences are shown in Q70466-68.
 CC Other specific peptides generated by these generic sequences are shown in
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
 CC comprising at least two functional regions - a binding domain with
 CC affinity for a ligand and a second effector peptide portion that is
 CC chemically or biologically active. They may further comprise a linker
 CC peptide between the 2 domains. The oligonucleotides are also designed so
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned
 CC in, or flanking, the unpredicted or variant residues. These residues
 CC confer some degree of conformational rigidity to the peptides. The TSARs
 CC deliver a chemically or biologically active moiety, eg. metal ion,
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
 CC cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need
 CC for complex methods of hybridoma formation or in vivo antibody
 CC production. The TSARs are easily characterised and have designed
 CC activity allowing direct and rapid detection in a screening process.
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;
 Query Match 1.3%; Score 38; DB 12; Length 114;
 Best Local Similarity 4.5%; Pred. No. 9.79e-07;
 Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;
 Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 52
 Cp 1527 GAAGCGGCTGCTGTGTCGAGCCAGCCAGAGCGCTGGGGCCACCGCGGCGAGCA 1468
 Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 114
 Cp 1467 GCACCGCAGGAGCGGTACACCTGCAGGGGCTGCTGTGCTGGCGGAGCAGC 1416
 RESULT 13
 ID Q70468 standard; DNA; 114 BP.
 AC Q70468;
 DE 05-APR-1995 (first entry)
 KW Generic DNA sequence to generate a random TSAR peptide library.
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; generic; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 55..60
 FT /tag= a
 FT /note= "this sequence represents 'Z'; 2 can be a
 FT sequence of 6, 9 or 12 nucleotides (see
 FT comments)"
 PN WO9418318-A.
 PD 18-AUG-1994.
 PE 01-FEB-1994; U00977.
 PR 01-FEB-1993; US-013416.
 PR 30-DEC-1993; US-176500.
 PR 31-JAN-1994; US-189331.
 PA (UUNC-) UNIV NORTH CAROLINA.
 PI Fowlkes DM, Kay BK;
 DR WPI: 94-279739/34.
 DR P-PSDB; R65154.
 DR Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PT comprising a binding domain and an effector domain
 PS Disclosure; Page 35; 255pp; English.
 CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be
 CC represented as follows: X(NNB)6(TGC)(NNB)12(NNB)14(TGC)(NNB)3Y. X
 CC and Y are flanking restriction sites (X is not the same as Y) that are
 CC not specified further. Other generic sequences are shown in Q70466-68.
 CC Other specific peptides generated by these generic sequences are shown in
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
 CC comprising at least two functional regions - a binding domain with
 CC affinity for a ligand and a second effector peptide portion that is
 CC chemically or biologically active. They may further comprise a linker
 CC peptide between the 2 domains. The oligonucleotides are also designed so
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned
 CC in, or flanking, the unpredicted or variant residues. These residues
 CC confer some degree of conformational rigidity to the peptides. The TSARs
 CC deliver a chemically or biologically active moiety, eg. metal ion,
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
 CC cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need
 CC for complex methods of hybridoma formation or in vivo antibody
 CC production. The TSARs are easily characterised and have designed
 CC activity allowing direct and rapid detection in a screening process.
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;
 Query Match 1.3%; Score 38; DB 12; Length 114;
 Best Local Similarity 4.5%; Pred. No. 9.79e-07;
 Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;
 Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 52
 Cp 1527 GAAGCGGCTGCTGTGTCGAGCCAGCCAGAGCGCTGGGGCCACCGCGGCGAGCA 1468
 Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 114
 Cp 1467 GCACCGCAGGAGCGGTACACCTGCAGGGGCTGCTGTGCTGGCGGAGCAGC 1416

FT /tag= a
 FT /note= "this sequence represents 'Z'; 2 can be a
 FT sequence of 6, 9 or 12 nucleotides (see
 FT comments)"
 PN WO9418318-A.
 PD 18-AUG-1994.
 PE 01-FEB-1994; U00977.
 PR 01-FEB-1993; US-013416.
 PR 30-DEC-1993; US-176500.
 PR 31-JAN-1994; US-189331.
 PA (UUNC-) UNIV NORTH CAROLINA.
 PI Fowlkes DM, Kay BK;
 DR WPI: 94-279739/34.
 DR P-PSDB; R65154.
 DR Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PT comprising a binding domain and an effector domain
 PS Disclosure; Page 35; 255pp; English.
 CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be
 CC represented as follows: X(NNB)11(TGC)(NNB)62(NNB)7(TGC)(NNB)10Y. X
 CC and Y are flanking restriction sites (X is not the same as Y) that are
 CC not specified further. Other generic sequences are shown in Q70466-68.
 CC Other specific peptides generated by these generic sequences are shown in
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
 CC comprising at least two functional regions - a binding domain with
 CC affinity for a ligand and a second effector peptide portion that is
 CC chemically or biologically active. They may further comprise a linker
 CC peptide between the 2 domains. The oligonucleotides are also designed so
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned
 CC in, or flanking, the unpredicted or variant residues. These residues
 CC confer some degree of conformational rigidity to the peptides. The TSARs
 CC deliver a chemically or biologically active moiety, eg. metal ion,
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
 CC cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need
 CC for complex methods of hybridoma formation or in vivo antibody
 CC production. The TSARs are easily characterised and have designed activity
 CC allowing direct and rapid detection in a screening process.
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;
 Query Match 1.3%; Score 38; DB 12; Length 114;
 Best Local Similarity 4.5%; Pred. No. 9.79e-07;
 Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;
 Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 62
 Cp 1542 CTTGGTGTCTGAGGAGCGGCTGCTGTGCTGAGCCCGAGAGCGGTACACCTGCAGGGGCTGTG 1431
 Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 114
 Cp 1482 CAGCGCGGCGAGCGCGCGCGAGCGGTACACCTGCAGGGGCTGTG 1431
 RESULT 14
 ID Q70467 standard; DNA; 114 BP.
 AC Q70467;
 DE 05-APR-1995 (first entry)
 KW Generic DNA sequence to generate a random TSAR peptide library.
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; generic; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 55..60
 FT /tag= a
 FT /note= "this sequence represents 'Z'; 2 can be a
 FT sequence of 6, 9 or 12 nucleotides (see
 FT comments)"
 PN WO9418318-A.
 PD 18-AUG-1994.
 PE 01-FEB-1994; U00977.

PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNK-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
DR P-PSDB: R55153.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)11Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R55151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.
CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.3%; Score 38; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 9,79e-07;
Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 62
Cp 1557 CAGGAGATGAACCTTGTGGTTCCTGAGGAAGCGGGTTCGTGTGCTGGAGGCCCA 1498
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnn 114
Cp 1497 GAGCGCTGGGGCCACGCGCGCGCAGCAGCGCCGCCGACGAGCGGTACACC 1446

RESULT 15
ID Q70470 standard; DNA; 114 BP.
AC Q70470;
DT 10-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "encoded by Z (see comments)"

PN W09418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNK-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
DR P-PSDB: R58378.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins

PT comprising a binding domain and an effector domain
PS Disclosure; Page 36; 255pp; English.
CC Q70470 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)4(CAC)(NNB)82(NNB)6(CAC)(NNB)8
CC -(CAC)2(NNB)Y. X and Y are flanking restriction sites (X is not the same
CC as Y) that are not specified further. The peptides generated by this and
CC other generic sequences (Q70471-73) have invariant histidine residues
CC incorporated into variant sequences. TSARs are concatenated
CC heterofunctional proteins or peptides, comprising at least two functional
CC regions - a binding domain with affinity for a ligand and a second
CC effector peptide portion that is chemically or biologically active. They
CC may further comprise a linker peptide between the 2 domains. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 5 A; 10 C; 0 G; 0 T;

Query Match 1.3%; Score 36; DB 12; Length 114;
Best Local Similarity 7.7%; Pred. No. 1.03e-05;
Matches 8; Conservative 30; Mismatches 66; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 62
Cp 2625 CAGCTTGTCTCCATGTCGCGGTAGCACAGGCTGCAGAGCAGCGTGAGAGATGGAGCC 2566
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnn 106
Cp 2565 CTGCGGGATCCCTGTGCACTGGAGCTGAGGACTTGCCTGTATGC 2522

Search completed: Fri Dec 25 15:11:06 1998
Job time : 1239 secs.

THIS PAGE BLANK (USPTO)

W P S R L A
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPSrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
 with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 15:16:55 1998; MasPar time 5202.31 seconds
 Tabular output not generated. 586.621 Million cell updates/sec

Title: >US-08-951-733-13
 Description: (1-2848) from US08951733.seq (1 of 3)
 Perfect Score: 2848
 N.A. Sequence: 1 CACGCGTCCGGCAGCGTG.....GATCGCGCCGACGGCCTAT 2848
 Comp: GTGGCAGGCGCGTGGCGAC.....CTACGCGCGGTGGCGGATA

Scoring table: TABLE default
 Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 1665728 seqs, 535777161 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: n-pending
 1:P9 2:U6000 3:U6001 4:U6002 5:U6003 6:U6004 7:U6005
 8:U6006 9:U6007 10:U6008 11:U7 12:U80 13:U81 14:U82
 15:U83 16:U84A 17:U84B 18:U85 19:U86 20:U87 21:U88
 22:U89 23:U90A 24:U90B 25:U91 26:NEWP 27:NEWU6 28:NEWU8
 29:NEWU9

Statistics: Mean 10.979; Variance 4.476; scale 2.453

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|-----------------------------------|-----------|
| 1 | 2848 | 100.0 | 2848 | 22 | US-08-951- Sequence 13, Applicati | 0.00e+00 |
| 2 | 2848 | 100.0 | 2798 | 22 | US-08-951- Sequence 19, Applicati | 0.00e+00 |
| 3 | 2840 | 99.7 | 7029 | 22 | US-08-911- Sequence 1, Applicati | 0.00e+00 |
| 4 | 2837 | 99.6 | 4015 | 21 | US-08-854- Sequence 224, Applicat | 0.00e+00 |
| 5 | 2837 | 99.6 | 4015 | 22 | US-08-912- Sequence 1, Applicati | 0.00e+00 |
| 6 | 2837 | 99.6 | 4015 | 24 | US-09-052- Sequence 1, Applicati | 0.00e+00 |
| 7 | 2834 | 99.5 | 4023 | 23 | US-09-026- Sequence 35, Applicati | 0.00e+00 |
| 8 | 2821 | 88.5 | 4029 | 21 | US-08-854- Sequence 173, Applicat | 0.00e+00 |
| 9 | 2521 | 88.5 | 4029 | 21 | US-08-851- Sequence 173, Applicat | 0.00e+00 |
| 10 | 2444 | 85.8 | 3855 | 22 | US-08-911- Sequence 18, Applicati | 0.00e+00 |
| 11 | 2444 | 85.8 | 3855 | 22 | US-08-912- Sequence 4, Applicatio | 0.00e+00 |
| 12 | 1497 | 52.6 | 4200 | 22 | US-08-912- Sequence 6, Applicatio | 0.00e+00 |
| 13 | 825 | 29.0 | 3496 | 23 | US-09-042- Sequence 1, Applicati | 0.00e+00 |
| 14 | 816 | 28.7 | 3346 | 23 | US-09-026- Sequence 36, Applicati | 0.00e+00 |
| 15 | 767 | 26.9 | 2176 | 22 | US-08-912- Sequence 3, Applicatio | 0.00e+00 |

16 735 25.8 2171 21 US-08-854- Sequence 100, Applicat 0.00e+00
 17 735 25.8 2171 21 US-08-846- Sequence 100, Applicat 0.00e+00
 18 735 25.8 2171 21 US-08-851- Sequence 100, Applicat 0.00e+00
 19 535 18.8 535 22 US-08-911- Sequence 52, Applicati 0.00e+00
 20 460 16.2 550 22 US-08-911- Sequence 54, Applicati 0.00e+00
 21 385 13.5 389 21 US-08-844- Sequence 62, Applicati 0.00e+00
 22 385 13.5 389 21 US-08-846- Sequence 62, Applicati 0.00e+00
 23 385 13.5 389 21 US-08-851- Sequence 62, Applicati 0.00e+00
 24 385 13.5 389 22 US-08-911- Sequence 17, Applicati 0.00e+00
 25 385 13.5 389 21 US-08-854- Sequence 62, Applicati 0.00e+00
 26 385 13.5 389 22 US-08-912- Sequence 8, Applicatio 0.00e+00
 27 182 6.4 182 22 US-08-912- Sequence 9, Applicatio 7.34e-146
 28 120 4.2 2651 23 US-09-042- Sequence 5, Applicatio 1.73e-84
 29 97 3.4 240 22 US-08-912- Sequence 7, Applicatio 2.26e-62
 30 69 2.4 477 11 US-07-904- Sequence 10, Applicati 2.34e-36
 31 69 2.4 477 11 US-07-904- Sequence 10, Applicati 2.34e-36
 32 68 2.4 578 18 US-08-569- Sequence 14, Applicati 1.87e-35
 33 68 2.4 578 19 US-08-624- Sequence 14, Applicati 1.87e-35
 34 68 2.4 578 19 US-08-624- Sequence 14, Applicati 1.87e-35
 35 68 2.4 578 18 US-08-569- Sequence 14, Applicati 1.87e-35
 36 69 2.4 1808 23 US-09-042- Sequence 4, Applicatio 2.34e-36
 37 67 2.4 7218 17 US-08-466- Sequence 14, Applicati 1.48e-34
 38 66 2.3 383 16 US-08-446- Sequence 8, Applicatio 1.17e-33
 39 66 2.3 383 15 US-08-311- Sequence 8, Applicatio 1.17e-33
 40 66 2.3 383 17 US-08-451- Sequence 8, Applicatio 1.17e-33
 41 66 2.3 383 16 US-08-446- Sequence 8, Applicatio 1.17e-33
 42 66 2.3 383 16 US-08-446- Sequence 8, Applicatio 1.17e-33
 43 66 2.3 383 16 US-08-446- Sequence 8, Applicatio 1.17e-33
 44 62 2.2 340 19 US-08-617- Sequence 9, Applicatio 4.27e-30
 45 62 2.2 383 17 US-08-451- Sequence 8, Applicatio 4.27e-30

ALIGNMENTS

RESULT 1
 ID US-08-951-733-13 STANDARD; DNA; UNC; 2848 BP.
 AC xxxxxx

Sequence 13, Application US/08951733
 Sequence 13, Application US/08951733

GENERAL INFORMATION:

APPLICANT: Harrington, Lea A.

APPLICANT: Robinson, Murray O.

TITLE OF INVENTION: NOVEL GENES ENCODING TELOWERASE PROTEINS

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Inc.

STREET: One Amgen Center Drive

CITY: Thousand Oaks

STATE: CA

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/951,733

FILING DATE: 16-OCT-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/873,039

FILING DATE: 11-JUN-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/751,189

FILING DATE: 15-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Oleski, Nancy A.

REGISTRATION NUMBER: 34,688

REFERENCE/DOCKET NUMBER: A-433B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (805) 447-6504

CC TELEFAX: (805) 499-8011
CC INFORMATION FOR SEQ ID NO: 13:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2848 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
SQ SEQUENCE 2848 BP; 437 A; 978 C; 945 G; 488 T; 0 OTHER.

Query Match 100.0%; Score 2848; DB 22; Length 2848;
Best Local Similarity 100.0%; Pred. No. 0.0e+00;
Matches 2848; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dbb 1 CACGGCTCCGGGAGCGCTGCTGCTGCGACAGTGGGAAGCCCTGGCCCCGGGCAAC 60
Qy 1 CACGGCTCCGGGAGCGCTGCGCTGCTGCGACAGTGGGAAGCCCTGGCCCCGGGCAAC 60

Dbb 61 CCGCGATCCGGGCGCTCCCGCTGCCAGCGCTGCGCTGCGCTGCGCGACCACTAC 120
Qy 61 CCGCGATCCGGGCGCTCCCGCTGCCAGCGCTGCGCTGCGCTGCGCGACCACTAC 120

Dbb 121 CCGAGAGTCTGCGCTGCGACAGTTCGTGCGCGCTGGGCGCCAGGCGTGGGCGTG 180
Qy 121 CCGAGAGTCTGCGCTGCGACAGTTCGTGCGCGCTGGGCGCCAGGCGTGGGCGTG 180

Dbb 181 GTGACGCGGGGACCCGGCGCTTTCGCGCGCTGGTGGCGCCAGTGCCTGTGCGTG 240
Qy 181 GTGACGCGGGGACCCGGCGCTTTCGCGCGCTGGTGGCGCCAGTGCCTGTGCGTG 240

Dbb 241 CCGTGGAGCGAGCG 300
Qy 241 CCGTGGAGCGAGCG 300

Dbb 301 GAGCTGTGCGCGAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGAGAGAGCTGTGGCC 360
Qy 301 GAGCTGTGCGCGAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGAGAGAGCTGTGGCC 360

Dbb 361 TTCGGCTTCGCGCTGCTGACAGGGGCGCGCGGGGCGCGCGCGCGCGCGCGCGCGCG 420
Qy 361 TTCGGCTTCGCGCTGCTGACAGGGGCGCGCGGGGCGCGCGCGCGCGCGCGCGCGCG 420

Dbb 421 GTGCGAGTACTTGTGCGCGAGCGGTGACGAGCGACTGCGGGGAGCGGGCGTGGGG 480
Qy 421 GTGCGAGTACTTGTGCGCGAGCGGTGACGAGCGACTGCGGGGAGCGGGCGTGGGG 480

Dbb 481 CTGCTCTGCGCGCGCTGGGCGACAGCTGCTGTTTCACTGCTGCGAGCTGCGCGCTC 540
Qy 481 CTGCTCTGCGCGCGCTGGGCGACAGCTGCTGTTTCACTGCTGCGAGCTGCGCGCTC 540

Dbb 541 TTTGTGCTGGTGGCTCCAGCTGCGGCTACAGGTTGCGGGCGCGCGCGCTGTACCACTC 600
Qy 541 TTTGTGCTGGTGGCTCCAGCTGCGGCTACAGGTTGCGGGCGCGCGCGCTGTACCACTC 600

Dbb 601 GCGCTGCCACTCAGCG 660
Qy 601 GCGCTGCCACTCAGCG 660

Dbb 661 TCGGAACGGGCTGGAACCATAGCTGACGAGCGCGCGGCTCCCTGCGCGCTGCCAGCC 720
Qy 661 TCGGAACGGGCTGGAACCATAGCTGACGAGCGCGCGGCTCCCTGCGCGCTGCCAGCC 720

Dbb 721 CCGGGTGCAGGAGCGCGGGGCGAGTGCAGCGGAAGTCTGCCGTGCGCGCGCGCGCGCG 780
Qy 721 CCGGGTGCAGGAGCGCGGGGCGAGTGCAGCGGAAGTCTGCCGTGCGCGCGCGCGCGCG 780

Dbb 781 AGCGTGGCGCTGCCCTCAGCGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
Qy 781 AGCGTGGCGCTGCCCTCAGCGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840

Dbb 841 CCGGGGAGGACCGTGGACCGAGTACCGTGTGTTTCTGTTGTTGTTGTTGTTGTTGTT 900
Qy 841 CCGGGGAGGACCGTGGACCGAGTACCGTGTGTTTCTGTTGTTGTTGTTGTTGTTGTT 900

Dbb 901 GCCGAAGAACCACTCTTTTGGAGGTCGCTCTCTGTGACGCGGCCACTTCCACCCATCC 960
Qy 901 GCCGAAGAACCACTCTTTTGGAGGTCGCTCTCTGTGACGCGGCCACTTCCACCCATCC 960

Dbb 961 GTGGGCGCGAGCACCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Qy 961 GTGGGCGCGAGCACCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020

Dbb 1021 ACGGCTTTCGCGCGGTCGTAGCGCGAGACCAAGCACTTCTCTACTCTCAGGGGACAAG 1080
Qy 1021 ACGGCTTTCGCGCGGTCGTAGCGCGAGACCAAGCACTTCTCTACTCTCAGGGGACAAG 1080

Dbb 1081 GAGCAGCTGCGGCGCTCTCTCTACTAGCTCTCTGAGGCGCGAGCTGAGCTGCGGCTCG 1140
Qy 1081 GAGCAGCTGCGGCGCTCTCTCTACTAGCTCTCTGAGGCGCGAGCTGAGCTGCGGCTCG 1140

Dbb 1141 AGGCTGCTGGAGACCATCTTCTGCTGCTGAGGCGCTTGGATGCCAGGAGCTTCCCGCAG 1200
Qy 1141 AGGCTGCTGGAGACCATCTTCTGCTGCTGAGGCGCTTGGATGCCAGGAGCTTCCCGCAG 1200

Dbb 1201 TTGCGCGCGCTGCCCGAGCGCTACTTGGCAATGGGCGCGCTGTTTCTGAGCTGCTGGG 1260
Qy 1201 TTGCGCGCGCTGCCCGAGCGCTACTTGGCAATGGGCGCGCTGTTTCTGAGCTGCTGGG 1260

Dbb 1261 AACCAACGCGAGTGCCTTCTCAAGAGCGACTGCCCGCTGCGAGCTGCG 1320
Qy 1261 AACCAACGCGAGTGCCTTCTCAAGAGCGACTGCCCGCTGCGAGCTGCG 1320

Dbb 1321 GTCACCCAGCAGCGGCTGCTGTCGCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Qy 1321 GTCACCCAGCAGCGGCTGCTGTCGCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCG 1380

Dbb 1381 GAGGAGGAGGACAGACCGCGCTGCTGCGCGAGTGCCTCGGCGAGCAGCAGCGCGCG 1440
Qy 1381 GAGGAGGAGGACAGACCGCGCTGCTGCGCGAGTGCCTCGGCGAGCAGCAGCGCGCG 1440

Dbb 1441 TGGCAGTGTACGGCTGCTGCGGCGCTGCTGCGCGCGCGCTGCTGCGCGCGCGCGCG 1500
Qy 1441 TGGCAGTGTACGGCTGCTGCGGCGCTGCTGCGCGCGCGCTGCTGCGCGCGCGCGCG 1500

Dbb 1501 GGCTCCAGCAGCAGCGCGCTTCTCAGGACACCAAGAGTTCATCTCCCTGGGG 1560
Qy 1501 GGCTCCAGCAGCAGCGCGCTTCTCAGGACACCAAGAGTTCATCTCCCTGGGG 1560

Dbb 1561 AAGCATGCCAGCTCTCGCTCAGGAGTGCCTGGAAGATGAGGTCGCGGAGTGCCT 1620
Qy 1561 AAGCATGCCAGCTCTCGCTCAGGAGTGCCTGGAAGATGAGGTCGCGGAGTGCCT 1620

Dbb 1621 TGGCTGCGAGGAGCGCGCGGTTGGCTGTTCGCGCGCGAGACCGCTTGGCTGAG 1680
Qy 1621 TGGCTGCGAGGAGCGCGCGGTTGGCTGTTCGCGCGCGAGACCGCTTGGCTGAG 1680

Dbb 1681 GAGATCTTGGCCAACTTCTCAGCTGCTGATGAGTGTGATGCTGCTGCTGCTGCTGCT 1740
Qy 1681 GAGATCTTGGCCAACTTCTCAGCTGCTGATGAGTGTGATGCTGCTGCTGCTGCTGCT 1740

Dbb 1741 TCTTCTTTTATGTACGAGACCACTTTCAGGAGACAGCTTCTTCTACCGGAG 1800
Qy 1741 TCTTCTTTTATGTACGAGACCACTTTCAGGAGACAGCTTCTTCTACCGGAG 1800

Dbb 1801 AGTGTCTGGAGCAAGTTGCAAGCATTCGAATCAGACGACTTGAAGAGGTTGAGCTG 1860
Qy 1801 AGTGTCTGGAGCAAGTTGCAAGCATTCGAATCAGACGACTTGAAGAGGTTGAGCTG 1860

Dbb 1861 CCGGAGCTGTCGGAAGCAGAGTCCAGGAGCATCGGAGCGCGCGCGCGCGCGCGCGCG 1920
Qy 1861 CCGGAGCTGTCGGAAGCAGAGTCCAGGAGCATCGGAGCGCGCGCGCGCGCGCGCGCG 1920

Dbb 1921 TCCAGACTCCGCTTCTCATCCCAAGCGCTGACGGCTGCGGGCTGCGGGCTGCGGGCTG 1980
Qy 1921 TCCAGACTCCGCTTCTCATCCCAAGCGCTGACGGCTGCGGGCTGCGGGCTGCGGGCTG 1980

| | | | |
|----|------|---|------|
| Db | 1981 | GTCTGGGAGCCAGAACGTTCCGCGAGAGAAAGAGGGCCGAGCGTCTCACTCTGAGGGTG | 2040 |
| Qy | 1981 | GTCTGGGAGCCAGAACGTTCCGCGAGAGAAAGAGGGCCGAGCGTCTCACTCTGAGGGTG | 2040 |
| Db | 2041 | AAGGCACTGTTTCAGCGTGCTCAACTACGAGCGGGCGGGCGCCCGCGCTCTCTGGGCGCC | 2100 |
| Qy | 2041 | AAGGCACTGTTTCAGCGTGCTCAACTACGAGCGGGCGGGCGCCCGCGCTCTCTGGGCGCC | 2100 |
| Db | 2101 | TCTGTGCTGGGCGCTGGACGATATCCACAGGGCGCTGGGCGACCTTTCGTGCTGCGTGTGCGG | 2160 |
| Qy | 2101 | TCTGTGCTGGGCGCTGGACGATATCCACAGGGCGCTGGGCGACCTTTCGTGCTGCGTGTGCGG | 2160 |
| Db | 2161 | GCCAGGACCCCGCGCTGAGCTGACTTTCATAGGTGGATGTGACGGGCGGTACGAC | 2220 |
| Qy | 2161 | GCCAGGACCCCGCGCTGAGCTGACTTTCATAGGTGGATGTGACGGGCGGTACGAC | 2220 |
| Db | 2221 | ACATCCCCCAGGACAGGCTCACGGAGGTCAATCGGCAGCATCATCAACCCCGAGAACG | 2280 |
| Qy | 2221 | ACATCCCCCAGGACAGGCTCACGGAGGTCAATCGGCAGCATCATCAACCCCGAGAACG | 2280 |
| Db | 2281 | TACTGGTGCCTGGTATGCCGTGTTCAGAAAGCCGCCATGGGCGACGTCCGGAAGGCC | 2340 |
| Qy | 2281 | TACTGGTGCCTGGTATGCCGTGTTCAGAAAGCCGCCATGGGCGACGTCCGGAAGGCC | 2340 |
| Db | 2341 | TTCAAGAGCCAGCTCTCACTTTCAGACACTCCAGCCGTACATGCGACAGTTTCGTGGCT | 2400 |
| Qy | 2341 | TTCAAGAGCCAGCTCTCACTTTCAGACACTCCAGCCGTACATGCGACAGTTTCGTGGCT | 2400 |
| Db | 2401 | CACCTCGAGGAGACAGCCCGCTGAGGATGCCGTCTCATCGAGCAGAGTCTCTCCCTG | 2460 |
| Qy | 2401 | CACCTCGAGGAGACAGCCCGCTGAGGATGCCGTCTCATCGAGCAGAGTCTCTCCCTG | 2460 |
| Db | 2461 | AATGAGCCAGAGTGGCTCTTCACGCTCTTCCTACGCTTCATGTGCCACGACGCGTG | 2520 |
| Qy | 2461 | AATGAGCCAGAGTGGCTCTTCACGCTCTTCCTACGCTTCATGTGCCACGACGCGTG | 2520 |
| Db | 2521 | CGCATCAGGGGCAAGTCTCTACGTCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCC | 2580 |
| Qy | 2521 | CGCATCAGGGGCAAGTCTCTACGTCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCC | 2580 |
| Db | 2581 | ACGCTGCTCTGACGCTGTGCTACGGGCACATGGAGACACAGCTGTTTGGGGGATTCGG | 2640 |
| Qy | 2581 | ACGCTGCTCTGACGCTGTGCTACGGGCACATGGAGAACAGCTGTTTGGGGGATTCGG | 2640 |
| Db | 2641 | CGGGACGGCGTCTCTCGGTTTGGTGGATGATTCTTTGTTGGTGACACCTCACTCACC | 2700 |
| Qy | 2641 | CGGGACGGCGTCTCTCGGTTTGGTGGATGATTCTTTGTTGGTGACACCTCACTCACC | 2700 |
| Db | 2701 | CAGCGAAACCTTCTCTCAGGACCTGTGTCGAGGTGTCCCTGAGTATGGCTGCGTGGT | 2760 |
| Qy | 2701 | CAGCGAAACCTTCTCTCAGGACCTGTGTCGAGGTGTCCCTGAGTATGGCTGCGTGGT | 2760 |
| Db | 2761 | AACCTTCGGAAGACAGTGGTGAACCTCCCTGTAGAGACGAGGCCCTGGGTGGCACGGCT | 2820 |
| Qy | 2761 | AACCTTCGGAAGACAGTGGTGAACCTCCCTGTAGAGACGAGGCCCTGGGTGGCACGGCT | 2820 |
| Db | 2821 | TTTGTTCAGATGCCGCCACCGCCCTAT | 2848 |
| Qy | 2821 | TTTGTTCAGATGCCGCCACCGCCCTAT | 2848 |

RESULT 2
ID US-08-951-733-19 STANDARD; DNA; UNC; 3798 BP.
AC xxxxxx

Sequence 19, Application US/08951733
Sequence 19, Application US/08951733
GENERAL INFORMATION:

CC GENERAL INFORMATION:
CC APPLICANT: Harrington, Lea A.
CC
CC APPLICANT: Robinson, Murray O.
CC
CC TITLE OF INVENTION: NOVEL GENETIC
CC
CC NUMBER OF SEQUENCES: 44
CC
CC CORRESPONDENCE ADDRESS:

| | | | |
|----|------|--|------|
| Db | 1141 | GACCATCTTTCTGGGTTTCAGGCGCTTCAGGATGCCAGGGACTCCCGCAGGTTGCCCGCGCT | 1200 |
| Qy | 1152 | GACCATCTTTCTGGGTTTCAGGCGCTTCAGGATGCCAGGGACTCCCGCAGGTTGCCCGCGCT | 1211 |
| Db | 1201 | GCCCCAGCGCTACTGCAAAATGGGCCCCCTGTTCTGAGAGTCTGTTGGGAACACAGCGCA | 1260 |
| Qy | 1212 | GCCCCAGCGCTACTGCAAAATGGGCCCCCTGTTCTGAGAGTCTGTTGGGAACACAGCGCA | 1271 |
| Db | 1261 | GTGCCCTACGGGTGCTCTCTCAAGACGCATGCCCGCTGCGAGCTGCGFTACCCCGACG | 1320 |
| Qy | 1272 | GTGCCCTACGGGTGCTCTCTCAAGACGCATGCCCGCTGCGAGCTGCGFTACCCCGACG | 1331 |
| Db | 1321 | AGCCGGTGTCTGTCCCGGGAAGCCCCAGGGCTCTGTGGCGCCCCCGCAGGAGGAGGA | 1380 |
| Qy | 1332 | AGCCGGTGTCTGTCCCGGGAAGCCCCAGGGCTCTGTGGCGCCCCCGCAGGAGGAGGA | 1391 |
| Db | 1381 | CACAGACCCCGTTCGGCTGTGTCAGCTGCTCCGCAGACACAGACGCCCTTGGCAGTGTGA | 1440 |
| Qy | 1392 | CACAGACCCCGTTCGGCTGTGTCAGCTGCTCCGCAGACACAGACGCCCTTGGCAGTGTGA | 1451 |
| Db | 1441 | CGGCTTCGTCGGGCTGCTCTCGCGCGCTGTGCCCGACCGCTCTGGGGCTCCAGGCA | 1500 |
| Qy | 1452 | CGGCTTCGTCGGGCTGCTCTCGCGCGCTGTGCCCGACCGCTCTGGGGCTCCAGGCA | 1511 |
| Db | 1501 | CAACGAACCGCGCTTCTCTCAGGAACACCAAGAAGTTCTCTCCCTGGGAAGCATGCCAA | 1560 |
| Qy | 1512 | CAACGAACCGCGCTTCTCTCAGGAACACCAAGAAGTTCTCTCCCTGGGAAGCATGCCAA | 1571 |
| Db | 1561 | GCTCTCGCTCAGAGAGTGACGTGGGAAGATGAGCTGCGGAGTGCGCTTGGCTGCGCAG | 1620 |
| Qy | 1572 | GCTCTCGCTCAGAGAGTGACGTGGGAAGATGAGCTGCGGAGTGCGCTTGGCTGCGCAG | 1631 |
| Db | 1621 | GAGCCAGGGGTTGGCTGTGTTCCGGCCGACAGACACCGTCTGGGTGAGGAGATCCTGGC | 1680 |
| Qy | 1632 | GAGCCAGGGGTTGGCTGTGTTCCGGCCGACAGACACCGTCTGGGTGAGGAGATCCTGGC | 1691 |
| Db | 1681 | CAAGTTCCTGCATGCGTGATGAGTGTACGTCTGCGAGCTGCTCAGGCTCTTCTCTTTTA | 1740 |
| Qy | 1692 | CAAGTTCCTGCATGCGTGATGAGTGTACGTCTGCGAGCTGCTCAGGCTCTTCTCTTTTA | 1751 |
| Db | 1741 | TGTCACGGAGACCACTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAG | 1800 |
| Qy | 1752 | TGTCACGGAGACCACTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAG | 1811 |
| Db | 1801 | CAAGTTCGAAGCATTTGGAAATCAGACAGACATTTGAAGAGGTGCAGCTGCGGAGCTGTC | 1860 |
| Qy | 1812 | CAAGTTCGAAGCATTTGGAAATCAGACAGACATTTGAAGAGGTGCAGCTGCGGAGCTGTC | 1871 |
| Db | 1861 | GGAAGCAGAGGTTCAGGCAGCATCGGAAGCCAGGCCCGCTGTGACGTCCAGACTCCG | 1920 |
| Qy | 1872 | GGAAGCAGAGGTTCAGGCAGCATCGGAAGCCAGGCCCGCTGTGACGTCCAGACTCCG | 1931 |
| Db | 1921 | CTTCATCCCCAAGCCTGACGGGCTCGGGCCGATTGTGAACATGGAATGAGTCTGGGAGC | 1980 |
| Qy | 1932 | CTTCATCCCCAAGCCTGACGGGCTCGGGCCGATTGTGAACATGGAATGAGTCTGGGAGC | 1991 |
| Db | 1981 | CAGAACGTTCCGCAGAGAAAGAGGGCCGAGCGTCTACCTCGAGGGTGAAGGCACTGTT | 2040 |
| Qy | 1992 | CAGAACGTTCCGCAGAGAAAGAGGGCCGAGCGTCTACCTCGAGGGTGAAGGCACTGTT | 2051 |
| Db | 2041 | CAGGCTGCTCAACTACGAGCGGGCGGGCCCGCGCTCTGGCGCTCTGTGCTGGG | 2100 |
| Qy | 2052 | CAGGCTGCTCAACTACGAGCGGGCGGGCCCGCGCTCTGGCGCTCTGTGCTGGG | 2111 |
| Db | 2101 | CCTGGAGCATATCCACAGGCGCTTGGCGCACTTCGTGCTGCTGTGGGGCCCGAGGACCC | 2160 |
| Qy | 2112 | CCTGGAGCATATCCACAGGCGCTTGGCGCACTTCGTGCTGCTGTGGGGCCCGAGGACCC | 2171 |
| Db | 2161 | GCGCCCTGAGCTGTACTTTGTCAAGGTGATGTGACGGGCGGTACGACACCATCCCCCA | 2220 |
| Qy | 2172 | GCGCCCTGAGCTGTACTTTGTCAAGGTGATGTGACGGGCGGTACGACACCATCCCCCA | 2231 |

| | | | |
|----|------|---|------|
| Db | 2221 | GGACAGGCTCACGAGGTGTAATGCCAGCAATCAATAAACCCAGCAACACGTACTCGGTGCG | 2280 |
| Qy | 2232 | GGACAGGCTCACGAGGTGTAATGCCAGCAATCAATAAACCCAGCAACACGTACTCGGTGCG | 2291 |
| Db | 2281 | TCGGTATGCCGTGGTCCAGAAGGCCCCATCGGCACGTCGCGAAGGCCTTCAAGAGCCA | 2340 |
| Qy | 2292 | TCGGTATGCCGTGGTCCAGAAGGCCCCATCGGCACGTCGCGAAGGCCTTCAAGAGCCA | 2351 |
| Db | 2341 | CGTCTCTACCTTGACAGACCTCCAGCCGTACATGGCAGACGTTGCTGGCTCACCTGCGAGGA | 2400 |
| Qy | 2352 | CGTCTCTACCTTGACAGACCTCCAGCCGTACATGGCAGACGTTGCTGGCTCACCTGCGAGGA | 2411 |
| Db | 2401 | GACACAGCCGCTGAGGGATGCCGTGCTCATGAGCAGAGCTCTCTCCCTGAATGAGGCCAG | 2460 |
| Qy | 2412 | GACACAGCCGCTGAGGGATGCCGTGCTCATGAGCAGAGCTCTCTCCCTGAATGAGGCCAG | 2471 |
| Db | 2461 | CAGTGGGCTCTTCGACGCTCTTCCATCGCTTCATGTGCCACACGCCGTGCGCATCAGGGG | 2520 |
| Qy | 2472 | CAGTGGGCTCTTCGACGCTCTTCCATCGCTTCATGTGCCACACGCCGTGCGCATCAGGGG | 2531 |
| Db | 2521 | CAAGTCTCTACGTCCAGTGCCAGGGGATCCCGCAGGGGCTCCATCTCTCCACGCTGCTCTG | 2580 |
| Qy | 2532 | CAAGTCTCTACGTCCAGTGCCAGGGGATCCCGCAGGGGCTCCATCTCTCTCCACGCTGCTCTG | 2591 |
| Db | 2581 | CAGCCTGTCTACGGCGACATGGAGAACAGCTGTTGCGGGGATTCGGCGGACGGGCT | 2640 |
| Qy | 2592 | CAGCCTGTCTACGGCGACATGGAGAACAGCTGTTGCGGGGATTCGGCGGACGGGCT | 2651 |
| Db | 2641 | GCTCCTCGCTTGGTGGATGATTCTTGTTGGTGACACCTCACCTCACCCACGCGAAAC | 2700 |
| Qy | 2652 | GCTCCTCGCTTGGTGGATGATTCTTGTTGGTGACACCTCACCTCACCCACGCGAAAC | 2711 |
| Db | 2701 | CTTCCTCAGAACCCCTGGTCGGAGGTGTCCTCAGTATGGCTGCGTGGTGAACCTTGGCGAA | 2760 |
| Qy | 2712 | CTTCCTCAGAACCCCTGGTCGGAGGTGTCCTCAGTATGGCTGCGTGGTGAACCTTGGCGAA | 2771 |
| Db | 2761 | GACAGTGGTGAACCTTCCTCTGTAGAAGCAGAGCCCTGGTGGCAGGGCTTTGTTTCAGAT | 2820 |
| Qy | 2772 | GACAGTGGTGAACCTTCCTCTGTAGAAGCAGAGCCCTGGTGGCAGGGCTTTGTTTCAGAT | 2831 |
| Db | 2821 | GCCGGGCCACGGCCTAT | 2837 |
| Qy | 2832 | GCCGGGCCACGGCCTAT | 2848 |

RESULT 5
ID US-08-912-951-1 STANDARD; DNA; UNC; 4015 BP.
AC xxxxxx

Sequence 1, Application US/08912951
Sequence 1, Application US/08912951
GENERAL INFORMATION:

| | |
|----|-------------------------------|
| CC | CONCURRENT INFORMATION. |
| CC | APPLICANT: Cech, Thomas R. |
| CC | APPLICANT: Lingner, Joachim |
| CC | APPLICANT: Nakamura, Toru |
| CC | APPLICANT: Chapman, Karen B. |
| CC | APPLICANT: Morin, Gregg B. |
| CC | APPLICANT: Harley, Calvin. |
| CC | APPLICANT: Andrews, William H |
| CC | TITLE OF INVENTION: HUMAN TEL |
| CC | TITLE OF INVENTION: THERAPEUT |
| CC | NUMBER OF SEQUENCES: 335 |

CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
City: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

QY 1392 CACAGACCCCGTCCGCTGGTGCAGCTGCTCGGCCAGACAGACGCCCTCGCAGGTGTA 1451
Db 1441 CGGCTTCGTGGGGCTGCTCGCGCCGCTGGTGGCCCCAGCCCTCTGGGGCTCCAGGCA 1500
QY 1452 CGGCTTCGTGGGGCTGCTCGCGCCGCTGGTGGCCCCAGCCCTCTGGGGCTCCAGGCA 1511
Db 1501 CAACGAAGCCGCTTCCTCAGAGACACCAAGAAGTTATCTCCCTGGGGAAGCATGCCAA 1560
QY 1512 CAACGAAGCCGCTTCCTCAGAGACACCAAGAAGTTATCTCCCTGGGGAAGCATGCCAA 1571
Db 1561 GCTCTCGCTCAGGAGCTGAGCTGGAAGATGAGCGTGGGGACTGCGCTTGGCTGCGGAG 1620
QY 1572 GCTCTCGCTCAGGAGCTGAGCTGGAAGATGAGCGTGGGGACTGCGCTTGGCTGCGGAG 1631
Db 1621 GAGCCAGGGTGGCTGTGTTCGGCGCCAGAGACCGCTGTGCTGAGGAGATCCTGGC 1680
QY 1632 GAGCCAGGGTGGCTGTGTTCGGCGCCAGAGACCGCTGTGCTGAGGAGATCCTGGC 1691
Db 1681 CAAGTTCCTGCACTGGCTGATGATGTGTAGTGTGCTGCTGAGCTGCTCAGGTCTTTT 1740
QY 1692 CAAGTTCCTGCACTGGCTGATGATGTGTAGTGTGCTGCTGAGCTGCTCAGGTCTTTT 1751
Db 1741 TGTCAAGGAGACCACTGTTTCAAAAGAACAGAGCTCTTTTCTACCGGAAGAGTGTGGAG 1800
QY 1752 TGTCAAGGAGACCACTGTTTCAAAAGAACAGAGCTCTTTTCTACCGGAAGAGTGTGGAG 1811
Db 1801 CAAGTTGCAAGCATTGAATCAGACAGCACTTGAAGAGGTGACGTGCGGGAGCTGTC 1860
QY 1812 CAAGTTGCAAGCATTGAATCAGACAGCACTTGAAGAGGTGACGTGCGGGAGCTGTC 1871
Db 1861 GGAAGCAGAGTCAGGAGCATCGGAGAGCCAGGCCCGCTGCTGAGCTCCAGACTCCG 1920
QY 1872 GGAAGCAGAGTCAGGAGCATCGGAGAGCCAGGCCCGCTGCTGAGCTCCAGACTCCG 1931
Db 1921 CTTATCCCAAGCTGACGGCTGCGCGCATTTGTGAACATGACTACGTCTGCGGAGC 1980
QY 1932 CTTATCCCAAGCTGACGGCTGCGCGCATTTGTGAACATGACTACGTCTGCGGAGC 1991
Db 1981 CAGAACGTTCCGAGAGAAAAGAGGGCGGAGCGTCTACCTCGAGGGTGAAGGCACTGTT 2040
QY 1992 CAGAACGTTCCGAGAGAAAAGAGGGCGGAGCGTCTACCTCGAGGGTGAAGGCACTGTT 2051
Db 2041 CAGCGTCTCAACTACGAGCGGGCGGCGCCCGCTGCTGAGCGCTCTGCTGCTGGG 2100
QY 2052 CAGCGTCTCAACTACGAGCGGGCGGCGCCCGCTGCTGAGCGCTCTGCTGCTGGG 2111
Db 2101 CCGTACGATATCACAGGGCTGCGCGCACCTTCTGCTGCTGCTGCGGGCCAGGAGCC 2160
QY 2112 CCGTACGATATCACAGGGCTGCGCGCACCTTCTGCTGCTGCTGCGGGCCAGGAGCC 2171
Db 2161 CCGCCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCCCCA 2220
QY 2172 CCGCCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCCCCA 2231
Db 2221 GGACAGGCTCAGGAGTCACTGCGAGCATCATCAAAACCCAGACAGCTACTCGGTGG 2280
QY 2232 GGACAGGCTCAGGAGTCACTGCGAGCATCATCAAAACCCAGACAGCTACTCGGTGG 2291
Db 2281 TCGGTATCGCTGCTGCAAGAGGGCGCCATGCGGACGCTCGCGAGGCTTTCAAGAGCCA 2340
QY 2292 TCGGTATCGCTGCTGCAAGAGGGCGCCATGCGGACGCTCGCGAGGCTTTCAAGAGCCA 2351
Db 2341 CGTCTACCTTGACAGACCTCCAGCCGCTACATGCGACAGTTCGTGGCTACCTGCGAGGA 2400
QY 2352 CGTCTACCTTGACAGACCTCCAGCCGCTACATGCGACAGTTCGTGGCTACCTGCGAGGA 2411
Db 2401 GACACGCCGCTGAGGATGCGCTGCTCATGAGCAGAGCTCTCTCCCTGATGAGGCCAG 2460
QY 2412 GACACGCCGCTGAGGATGCGCTGCTCATGAGCAGAGCTCTCTCCCTGATGAGGCCAG 2471
Db 2461 CAGTGGCTCTTCGACGCTTCTCCTACGCTTCTATGTCACACACCGCTGCGCATCAGGG 2520
QY 2472 CAGTGGCTCTTCGACGCTTCTCCTACGCTTCTATGTCACACACCGCTGCGCATCAGGG 2531

RESULT 6

ID US-09-052-919-1 STANDARD; DNA; UNC; 4015 BP.

AC xxxxxx

DT

DE Sequence 1, Application US/09052919

CC Sequence 1, Application US/09052919

CC GENERAL INFORMATION:

CC APPLICANT: Cech, Thomas R.

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin B.

CC APPLICANT: Andrews, William H.

CC TITLE OF INVENTION: Antisense Compositions for Detecting and

CC TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase

CC NUMBER OF SEQUENCES: 72

CC CORRESPONDENCE ADDRESS:

CC ADDRESS: Townsend and Townsend and Crew LLP

CC STREET: Two Embarcadero Center, Eighth Floor

CC CITY: San Francisco

CC STATE: California

CC COUNTRY: USA

CC ZIP: 94111-3834

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/09/052,919

CC FILING DATE: 31-MAR-1998

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/724,643

CC FILING DATE: 01-OCT-1996

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/844,419

CC FILING DATE: 18-APR-1997

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/846,017

CC FILING DATE: 25-APR-1997

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/851,843

CC FILING DATE: 06-MAY-1997

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/854,050

Db 1501 CAACGAACGCCGCTTCCCTCAGGAACACCAAGAAGTTTCATCTCCCTGGGGAAGCATGCCAA 1560
QY 1512 CAACGAACGCCGCTTCCCTCAGGAACACCAAGAAGTTTCATCTCCCTGGGGAAGCATGCCAA 1571
Db 1561 GCTCTCGCTGACGAGGCTGACGTGGAAGATGACGTGCGGGAAGTCCGCTTGGCTCGGAG 1620
QY 1572 GCTCTCGCTGACGAGGCTGACGTGGAAGATGACGTGCGGGAAGTCCGCTTGGCTCGGAG 1631
Db 1621 GAGCCAGGGGTGGCTGTGTTCCGCGCCGACAGACACCGTCTGCTGAGGAGATCCTGGC 1680
QY 1632 GAGCCAGGGGTGGCTGTGTTCCGCGCCGACAGACACCGTCTGCTGAGGAGATCCTGGC 1691
Db 1681 CAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCTGAGCTGCTCAGTCTTCTTTTA 1740
QY 1692 CAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCTGAGCTGCTCAGTCTTCTTTTA 1751
Db 1741 TGTCAGGAGACCGTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAG 1800
QY 1752 TGTCAGGAGACCGTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAG 1811
Db 1801 CAAGTTGCAAGCATTTGGAATCAGACAGCAGTGTGAAGAGGTCAGCTGCGGGAGCTGTC 1860
QY 1812 CAAGTTGCAAGCATTTGGAATCAGACAGCAGTGTGAAGAGGTCAGCTGCGGGAGCTGTC 1871
Db 1861 GGAAGCAGAGGTCAGGAGCATCGGGAAGCCAGGCCCGCCCTGCTGAGCTCCAGACTCCG 1920
QY 1872 GGAAGCAGAGGTCAGGAGCATCGGGAAGCCAGGCCCGCCCTGCTGAGCTCCAGACTCCG 1931
Db 1921 CTTTCATCCCAAGCTGAGCGGCTGGGCCGATTGTGAACATGGACTAGCTGCTGGAGC 1980
QY 1932 CTTTCATCCCAAGCTGAGCGGCTGGGCCGATTGTGAACATGGACTAGCTGCTGGAGC 1991
Db 1981 CAGAAGCTTCCGACAGAAAGAGGCGCAGGCTGTCACTCGAGGTTGAAGCACTGTT 2040
QY 1992 CAGAAGCTTCCGACAGAAAGAGGCGCAGGCTGTCACTCGAGGTTGAAGCACTGTT 2051
Db 2041 CAGCGTGTCAACTAGAGCGGCGCGCCCGCCCTGCTGGGCGCTCTGTGCTGGG 2100
QY 2052 CAGCGTGTCAACTAGAGCGGCGCGCCCGCCCTGCTGGGCGCTCTGTGCTGGG 2111
Db 2101 CTTGACGATATCCACAGGCTGCGGACCTTCGTGCTGCGTGTGCGGCGCCAGGACC 2160
QY 2112 CTTGACGATATCCACAGGCTGCGGACCTTCGTGCTGCGTGTGCGGCGCCAGGACC 2171
Db 2161 GCGCGCTGAGCTGTACTTGTCAAGTGTGATGTGACGGGCGGTACGACACCATCCCCA 2220
QY 2172 GCGCGCTGAGCTGTACTTGTCAAGTGTGATGTGACGGGCGGTACGACACCATCCCCA 2231
Db 2221 GGACAGGCTCAGGAGGTCATCGCCAGCATCATCAAAACCCAGAACACGCTACTGCGTGG 2280
QY 2232 GGACAGGCTCAGGAGGTCATCGCCAGCATCATCAAAACCCAGAACACGCTACTGCGTGG 2291
Db 2281 TCGGTATGCGGTGGTCCAGAGGCGCCCATGGGACGTCGCGAGGCTTCAAGAGCCA 2340
QY 2292 TCGGTATGCGGTGGTCCAGAGGCGCCCATGGGACGTCGCGAGGCTTCAAGAGCCA 2351
Db 2341 CGTCTCTACTTGCACAGACTCCAGCGCTACATGCGACAGTTCGTGGCTACCTCGAGGA 2400
QY 2352 CGTCTCTACTTGCACAGACTCCAGCGCTACATGCGACAGTTCGTGGCTACCTCGAGGA 2411
Db 2401 GACCAGCGCTGAGGATGCGGCTGCTCATCGACGAGAGCTCCTCCCTGAATGAGGCCAG 2460
QY 2412 GACCAGCGCTGAGGATGCGGCTGCTCATCGACGAGAGCTCCTCCCTGAATGAGGCCAG 2471
Db 2461 CAGTGGCTTTCGAGCTTCTTACGCTTCATGTGCCACCGCGCTGCGCATCAGGG 2520
QY 2472 CAGTGGCTTTCGAGCTTCTTACGCTTCATGTGCCACCGCGCTGCGCATCAGGG 2531
Db 2521 CAAGTCTCAGTCCAGTGCAGGGATCCCGAGGCTCCATCCTCTCCACGCTGCTCTG 2580
QY 2532 CAAGTCTCAGTCCAGTGCAGGGATCCCGAGGCTCCATCCTCTCCACGCTGCTCTG 2591

Db 2581 CAGCCTGTCTACGCGCACATGGAGAACAAAGCTGTTTGGGGGATTCGGCGGGACGGCT 2640
QY 2592 CAGCCTGTCTACGCGCACATGGAGAACAAAGCTGTTTGGGGGATTCGGCGGGACGGCT 2651
Db 2641 GCTCCTGCTTGGTGTGATGATTTCTTGTGTGACACCTCACCTCACCCACGCGAAAC 2700
QY 2652 GCTCCTGCTTGGTGTGATGATTTCTTGTGTGACACCTCACCTCACCCACGCGAAAC 2711
Db 2701 CTTCTCAGGACCTGTGTCGAGGTGTCCTGTAGATAGTGGTGTGCTGTAACCTTTCGCGAA 2760
QY 2712 CTTCTCAGGACCTGTGTCGAGGTGTCCTGTAGATAGTGGTGTGCTGTAACCTTTCGCGAA 2771
Db 2761 GACAGTGGTGAACCTTCCCTGTAGAGAGGAGGCGCTGGTGGCACGCTTTTGTTCAGAT 2820
QY 2772 GACAGTGGTGAACCTTCCCTGTAGAGAGGAGGCGCTGGTGGCACGCTTTTGTTCAGAT 2831
Db 2821 GCCGGCCACGCGCTAT 2837
QY 2832 GCCGGCCACGCGCTAT 2848
RESULT 7
ID US-09-026-981-35 STANDARD; DNA; UNC; 4023 BP.
AC xxxxx
DT Sequence 35, Application US/09026981
DE Sequence 35, Application US/09026981
CC GENERAL INFORMATION:
CC APPLICANT: Counter, Christopher M.
CC APPLICANT: Meyerson, Matthew
CC APPLICANT: Weinberg, Robert A.
CC TITLE OF INVENTION: Telomerase Catalytic Subunit Gene and
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Millitia Drive
CC CITY: Lexington
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/026,981
CC FILING DATE: 20-FEB-1998
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/064,322
CC FILING DATE: 30-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/055,762
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/054,549
CC FILING DATE: 01-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/047,151
CC FILING DATE: 20-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/038,750
CC FILING DATE: 20-FEB-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WHI97-11p4AM
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 781-861-6240
CC TELEFAX: 781-861-9540
CC INFORMATION FOR SEQ ID NO: 35:
CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 4023 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
SQ SEQUENCE 4023 BP; 668 A; 1363 C; 1277 G; 715 T; 0 OTHER.

Query Match 99.5%; Score 2834; DB 23; Length 4023;
Best Local Similarity 99.98; Pred. No. 0.00e+00;
Matches 2836; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 3 GGCAGCGCTCGCTCTGCTGCGACAGTGGGAAGCCCTGGCCCCCGCCACACCCCGCGATGC 62
QY |||||
Db 11 GGCAGCGCTCGCTCTGCTGCGACAGTGGGAAGCCCTGGCCCCCGCCACACCCCGCGATGC 70
QY |||||
Db 63 GCGCGCTCCCGTCCCGAGCGTGGCTCCCTGCTGCGAGCACTACCGAGAGTGC 122
QY |||||
Db 71 GCGCGCTCCCGTCCCGAGCGTGGCTCCCTGCTGCGAGCACTACCGAGAGTGC 130
QY |||||
Db 123 TGCGCTGGCCACGTTCTGCGCGCTGGGGCCCGCCAGGCTGGCGCTGGTCAGCGCG 182
QY |||||
Db 131 TGCGCTGGCCACGTTCTGCGCGCTGGGGCCCGCCAGGCTGGCGCTGGTCAGCGCG 190
QY |||||
Db 183 GGGACCGCGCGCTTTCCGCGCTGCTGGCCCAAGTGCCTGGTGGCTGGCCCTGGGACG 242
QY |||||
Db 191 GGGACCGCGCGCTTTCCGCGCTGCTGGCCCAAGTGCCTGGTGGCTGGCCCTGGGACG 250
QY |||||
Db 243 CAGCGCGCGCGCGCGCGCGCTTCCTTCGCGCAAGTGTCTGCTGAAGAGAGTGTGG 302
QY |||||
Db 251 CAGCGCGCGCGCGCGCGCGCTTCCTTCGCGCAAGTGTCTGCTGAAGAGAGTGTGG 310
QY |||||
Db 303 CCGAGTGTGCAGAGAGTGTGCGAGCGCGCGCGCAAGAGTGTCTGGCTTCGGCTTCG 362
QY |||||
Db 311 CCGAGTGTGCAGAGAGTGTGCGAGCGCGCGCGCAAGAGTGTCTGGCTTCGGCTTCG 370
QY |||||
Db 363 CGCTGTGTGAGCGCGCGCGCGCGCGCGCTTCACCAAGAGTGTGCGCGAGCT 422
QY |||||
Db 371 CGCTGTGTGAGCGCGCGCGCGCGCGCGCTTCACCAAGAGTGTGCGCGAGCT 430
QY |||||
Db 423 ACCTGCCCAACACGGTGACCGACGACTGCGGGGAGCGGGGCGTGGGGTGTGTTC 482
QY |||||
Db 431 ACCTGCCCAACACGGTGACCGACGACTGCGGGGAGCGGGGCGTGGGGTGTGTTC 490
QY |||||
Db 483 GCGCGTGGGCGAGAGTGTGCTGCTTACCTGTGCGACGCTGGCGCTCTTTGTGTGG 542
QY |||||
Db 491 GCGCGTGGGCGAGAGTGTGCTGCTTACCTGTGCGACGCTGGCGCTCTTTGTGTGG 550
QY |||||
Db 543 TGGTCTCCAGTGGCTTACCAGTGTGCGGGCGCGCTGTACAGTGGCGCTGCCA 602
QY |||||
Db 551 TGGTCTCCAGTGGCTTACCAGTGTGCGGGCGCGCTGTACAGTGGCGCTGCCA 610
QY |||||
Db 603 CTCAGGCCCGCGCGCGCGCGCTAGTGGACCCCGAGAGCGTCTGGGATGCGAAGCGG 662
QY |||||
Db 611 CTCAGGCCCGCGCGCGCGCGCTAGTGGACCCCGAGAGCGTCTGGGATGCGAAGCGG 670
QY |||||
Db 663 CCTGGACCACTAGCTCAGGAGCGCGGCTCCCGCTGGCGCTGCCAGCCCGGGTGGCA 722
QY |||||
Db 671 CCTGGACCACTAGCTCAGGAGCGCGGCTCCCGCTGGCGCTGCCAGCCCGGGTGGCA 730
QY |||||
Db 723 GGAGCGCGGGGCGAGTGCACAGCGAAGTCTGCGGTTGCCAAGAGCGCCAGGCGTGGCG 782
QY |||||
Db 731 GGAGCGCGGGGCGAGTGCACAGCGAAGTCTGCGGTTGCCAAGAGCGCCAGGCGTGGCG 790
QY |||||
Db 783 CTGCCCCGTAGCCGAGCGGACGCCGTTGGGCAAGGGTCTTGGGCCACCCGGCGCAGGA 842
QY |||||
Db 791 CTGCCCCGTAGCCGAGCGGACGCCGTTGGGCAAGGGTCTTGGGCCACCCGGCGCAGGA 850
QY |||||
Db 843 GCGGTGACCGAGTGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
QY |||||
Db 851 GCGGTGACCGAGTGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 910
QY |||||
Db 903 CCACCTCTTTGGAGGTCGCTCTCTGCGAGCGCGCGCTCCACCCCTGCTGCTGCTGCTGCT 962
QY |||||
Db 911 CCACCTCTTTGGAGGTCGCTCTCTGCGAGCGCGCGCTCCACCCCTGCTGCTGCTGCTGCT 970
QY |||||

Db 963 AGCACACGCGGCG 1022
QY |||||
Db 971 AGCACACGCGGCG 1030
QY |||||
Db 1023 CCCGGGTAGCGCGGAGACCAAGCACTTCTCTACTCTCAGGCGCAAGAGCAGCTGC 1082
QY |||||
Db 1031 CCCGGGTAGCGCGGAGACCAAGCACTTCTCTACTCTCAGGCGCAAGAGCAGCTGC 1090
QY |||||
Db 1083 GGGCTCTCTTCTACTACTAGCTCTCTGAGGCCAGCCCTGACTGGCGCTCGGAGGCTCGTGG 1142
QY |||||
Db 1091 GGGCTCTCTTCTACTACTAGCTCTCTGAGGCCAGCCCTGACTGGCGCTCGGAGGCTCGTGG 1150
QY |||||
Db 1143 AGACCATCTTCTGGGTTCCAGGCCCTGGATGCCAGGCACTCCCCGAGGTGGCCCCGCGC 1202
QY |||||
Db 1151 AGACCATCTTCTGGGTTCCAGGCCCTGGATGCCAGGCACTCCCCGAGGTGGCCCCGCGC 1210
QY |||||
Db 1203 TGCCCCACGCGCTACTGGCAAAATGCGGCCCTGTTCTTGGAGCTGCTTGGGAACACACGCG 1262
QY |||||
Db 1211 TGCCCCAGCGCTACTGGCAAAATGCGGCCCTGTTCTTGGAGCTGCTTGGGAACACACGCG 1270
QY |||||
Db 1263 AGTCCCCCTACGGGGTGTCTCTCAAGACGCACTGCCCGCTGCGAGCTGCGGTTCACCCAG 1322
QY |||||
Db 1271 AGTCCCCCTACGGGGTGTCTCTCAAGACGCACTGCCCGCTGCGAGCTGCGGTTCACCCAG 1330
QY |||||
Db 1323 CAGCGGTGTCTGTGCCGGGAGAACCCAGGCTCTGTGGCGGCCCGCCGAGGAGG 1382
QY |||||
Db 1331 CAGCGGTGTCTGTGCCGGGAGAACCCAGGCTCTGTGGCGGCCCGCCGAGGAGG 1390
QY |||||
Db 1383 ACACAGACCCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1442
QY |||||
Db 1391 ACACAGACCCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1450
QY |||||
Db 1443 ACAGCTTCGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1502
QY |||||
Db 1451 ACAGCTTCGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1510
QY |||||
Db 1503 ACACAGAACCCCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAAGCATGCCA 1562
QY |||||
Db 1511 ACACAGAACCCCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAAGCATGCCA 1570
QY |||||
Db 1563 AGCTCTGCTGCGAGGAGTGCAGTGGAGATGACGCTGCGGGGCTGCGCTGCTGCTGCTGCTGCTGCT 1622
QY |||||
Db 1571 AGCTCTGCTGCGAGGAGTGCAGTGGAGATGACGCTGCGGGGCTGCGCTGCTGCTGCTGCTGCT 1630
QY |||||
Db 1623 GGAGCCGAGGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1682
QY |||||
Db 1631 GGAGCCGAGGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1690
QY |||||
Db 1683 CCAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1742
QY |||||
Db 1691 CCAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1750
QY |||||
Db 1743 ATGTCAGGAGACCAAGCTTCAAGAGAACAGCTCTTTTCTACCGGAAGAGTGTCTGGA 1802
QY |||||
Db 1751 ATGTCAGGAGACCAAGCTTCAAGAGAACAGCTCTTTTCTACCGGAAGAGTGTCTGGA 1810
QY |||||
Db 1803 GCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGTCAGCTGCGGAGCTGT 1862
QY |||||
Db 1811 GCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGTCAGCTGCGGAGCTGT 1870
QY |||||
Db 1863 CGGAAGCAGAGGTCAGGACATCGGGAAGCCAGGCCCGCCCTGCTGACGCTCCAGACTCC 1922
QY |||||
Db 1871 CGGAAGCAGAGGTCAGGACATCGGGAAGCCAGGCCCGCCCTGCTGACGCTCCAGACTCC 1930
QY |||||
Db 1923 GCTTCATCCCAAGCTGACGGCTGCGGCCGATGTTGAACATGAGTACGTCGTGGAG 1982
QY |||||
Db 1931 GCTTCATCCCAAGCTGACGGCTGCGGCCGATGTTGAACATGAGTACGTCGTGGAG 1990
QY |||||
Db 1983 CCAGAACGTTCCGAGAGAAAGAGGCGCGAGCTCTCACCTCGAGGTTGAAGCACTGT 2042
QY |||||
Db 1991 CCAGAACGTTCCGAGAGAAAGAGGCGCGAGCTCTCACCTCGAGGTTGAAGCACTGT 2050
QY |||||

| | | | |
|--------|--|---|------|
| QY | 2411 | AGACGAGCCGCTGAGGATGCCGTGTCATCGACGAGCTCCTCCTGAATGAGCCA | 2470 |
| Db | 2455 | GCAGTGGCCTCTTCGACGCTTTCTACGCTTCATGTGCCACGCGGTGCGCATCAGG | 2514 |
| QY | 2471 | GCAGTGGCCTCTTCGACGCTTTCTACGCTTCATGTGCCACGCGGTGCGCATCAGG | 2530 |
| Db | 2515 | GCAAGTCTACGTCAGTCCAGTCCAGAGGATCCCGCAGGGTCCATCCTCTCCACGCTGCTCT | 2574 |
| QY | 2531 | GCAAGTCTACGTCAGTCCAGTCCAGAGGATCCCGCAGGGTCCATCCTCTCCACGCTGCTCT | 2590 |
| Db | 2575 | GCAGCCTGTGCTACGCGCATGGAGCAAGCTGTTTGGCGGGATTCGGCGGACGGGC | 2634 |
| QY | 2591 | GCAGCCTGTGCTACGCGCATGGAGCAAGCTGTTTGGCGGGATTCGGCGGACGGGC | 2650 |
| Db | 2635 | TGCTCTCGCCTTTGGTGGATGATTTCTTGTGGTGACACCTCAGCTCAGCCACGCAAAA | 2694 |
| QY | 2651 | TGCTCTCGCCTTTGGTGGATGATTTCTTGTGGTGACACCTCAGCTCAGCCACGCAAAA | 2710 |
| Db | 2695 | CTTCTCAGGACCCGTCGAGTGTCCTTGAGTATGGCTGGTGGTGAACCTTGTTCAGA | 2754 |
| QY | 2711 | CTTCTCAGGACCCGTCGAGTGTCCTTGAGTATGGCTGGTGGTGAACCTTGTTCAGA | 2770 |
| Db | 2755 | AGACACTGTGTAACCTCCCTGTAGAGAGCAGGAGCCCTGGGTGGCACGGCTTGTTCAGA | 2814 |
| QY | 2771 | AGACACTGTGTAACCTCCCTGTAGAGAGCAGGAGCCCTGGGTGGCACGGCTTGTTCAGA | 2830 |
| Db | 2815 | TGCGCGCCACCGGCTAT | 2832 |
| QY | 2831 | TGCGCGCCACCGGCTAT | 2848 |
| RESULT | 9 | | |
| ID | US-08-851-843-173 | STANDARD; DNA; UNC; 4029 BP. | |
| AC | xxxxxx | | |
| DT | | | |
| DE | Sequence 173, Application US/08851843 | | |
| CC | Sequence 173, Application US/08851843 | | |
| CC | GENERAL INFORMATION: | | |
| CC | APPLICANT: Cech, Thomas R. | | |
| CC | APPLICANT: Lingner, Joachim | | |
| CC | APPLICANT: Nakamura, Toru | | |
| CC | APPLICANT: Chapman, Karen B. | | |
| CC | APPLICANT: Morin, Gregg B. | | |
| CC | APPLICANT: Harley, Calvin | | |
| CC | APPLICANT: Andrews, William H. | | |
| CC | TITLE OF INVENTION: Novel Telomerase | | |
| CC | NUMBER OF SEQUENCES: 223 | | |
| CC | CORRESPONDENCE ADDRESS: | | |
| CC | ADDRESSEE: Townsend and Townsend and Crew LLP | | |
| CC | STREET: Two Embarcadero Center, 8th Floor | | |
| CC | CITY: San Francisco | | |
| CC | STATE: California | | |
| CC | COUNTRY: United States of America | | |
| CC | ZIP: 94111 | | |
| CC | COMPUTER READABLE FORM: | | |
| CC | MEDIUM TYPE: Floppy disk | | |
| CC | COMPUTER: IBM PC compatible | | |
| CC | OPERATING SYSTEM: PC-DOS/MS-DOS | | |
| CC | SOFTWARE: PatentIn Release #1.0, Version #1.30 | | |
| CC | CURRENT APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: US/08/851,843 | | |
| CC | FILING DATE: 06-MAY-1997 | | |
| CC | CLASSIFICATION: 536 | | |
| CC | PRIOR APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: US 08/846,017 | | |
| CC | FILING DATE: 25-APR-1997 | | |
| CC | CLASSIFICATION: 536 | | |
| CC | PRIOR APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: US 08/844,419 | | |
| CC | FILING DATE: 18-APR-1997 | | |
| CC | CLASSIFICATION: 536 | | |
| CC | PRIOR APPLICATION DATA: | | |

Db 659 CTGGAACCATAGCGTCAGGAGGCGGGGTCCCTCGGC-TGCCAGCCCGGGTGCAG 717
QY 672 CTGGAACCATAGCGTCAGGAGGCGGGGTCCCTCGGC-TGCCAGCCCGGGTGCAG 731
Db 718 GAGCGCGGGGAGTGCACAGCCGAAGTCTCGCGTTGCCCAAGAGGCCAGGGGTGCGC 777
QY 732 GAGCGCGGGGAGTGCACAGCCGAAGTCTCGCGTTGCCCAAGAGGCCAGGGGTGCGC 791
Db 778 TGCCCTTGACCGGAGCGGAGCGCCGTTGGCGAGGGTCTCTGGGCCACCCCGGCGAGGAC 837
QY 792 TGCCCTTGACCGGAGCGGAGCGCCGTTGGCGAGGGTCTCTGGGCCACCCCGGCGAGGAC 851
Db 838 GCCTGGACCGAGTGACCGTGGTTTCTGTGGTGTCACTGCCAGAGCCCGGAGAGC 897
QY 852 GCGTGGACCGAGTGACCGTGGTTTCTGTGGTGTCACTGCCAGAGCCCGGAGAGC 911
Db 898 CACCTCTTTGGAGGGTGCCTCTCTGCGACGCGCACTCCACCCATCCGTGGGCCGCCA 957
QY 912 CACCTCTTTGGAGGGTGCCTCTCTGCGACGCGCACTCCACCCATCCGTGGGCCGCCA 971
Db 958 GCACACGCGGGCCCCCATCCACATCGCGGCCACACAGTCC-TGGACACGCGTTGTCC 1016
QY 972 GCACACGCGGGCCCCCATCCACATCGCGGCCACACAGTCC-TGGACACGCGTTGTCC 1031
Db 1017 CCGGTGTACGCGGAGACCAAGCACTTCTCTACTCTCAGGCGACAAGNA-CACTCGGN 1075
QY 1032 CCGGTGTACGCGGAGACCAAGCACTTCTCTACTCTCAGGCGACAAGGAGCACTGCG 1091
Db 1076 CCCTCTCTCTACTCAATATATCTGAGGCCAGCCTCACTTGGCGTTCGGAGGTTCTGG 1135
QY 1092 GCCCTCTCTCTACTCAAGTCTCTGAGGCCAGCCTCACTTGGCGTTCGGAGGTTCTGGA 1151
Db 1136 AGACANTCTTTCTGGTTTCCAGGCTTGGATGCCAG-ATTCCCGCGAGGTTGCCGCCCT 1194
QY 1152 GACCATCTTTCTGGTTTCCAGGCTTGGATGCCAGGACTCCCGCAGGTTGCCGCCCT 1211
Db 1195 GCGCCAGGNTACTGGCAATCGGCCCTCTTCTGGAGCTGCTTGGGAACACGCGCA 1254
QY 1212 GCGCCAGGNTACTGGCAATCGGCCCTCTTCTGGAGCTGCTTGGGAACACGCGCA 1271
Db 1255 GTGCCCTACGGGTGTCTCTCAAGAGCGCACTGCCCGCTCGAGCTCGCGTCACTCCAGC 1314
QY 1272 GTGCCCTACGGGTGTCTCTCAAGAGCGCACTGCCCGCTCGAGCTCGCGTCACTCCAGC 1331
Db 1315 AGCGGTGTCTGTCCCGGAGAACCCAGGGCTCTGTGGCGGCCCGCCAGGAGGAGGA 1374
QY 1332 AGCGGTGTCTGTCCCGGAGAACCCAGGGCTCTGTGGCGGCCCGCCAGGAGGAGGA 1391
Db 1375 ACACAGACCCCGTCCGCTGTGTGAGCTGTCTCCCGCAGCAGCAGGCCCCCTGGCAGGTGT 1434
QY 1392 -CACAGACCCCGTCCGCTGTGTGAGCTGTCTCCCGCAGCAGCAGGCCCCCTGGCAGGTGT 1450
Db 1435 ACGGCTTCGTGCGGGCTGCTGCGCGCGGTGTGTGCCCGCGGCTCTGGGGCTCCAGGC 1494
QY 1451 ACGGCTTCGTGCGGGCTGCTGCGCGCGGTGTGTGCCCGCGGCTCTGGGGCTCCAGGC 1510
Db 1495 ACAACGAACCGCGCTTCTCAGGAACACCAAGAGTTTCACTTCCTCGGGAAGCATGCCA 1554
QY 1511 ACAACGAACCGCGCTTCTCAGGAACACCAAGAGTTTCACTTCCTCGGGAAGCATGCCA 1570
Db 1555 AGCTCTGCTGCAGGAGCTGACGTGGAGATGACGTGCGGGAGTGGCGTGGCTGCGCA 1614
QY 1571 AGCTCTGCTGCAGGAGCTGACGTGGAGATGACGTGCGGGAGTGGCGTGGCTGCGCA 1630
Db 1615 GGAGCCAGGGTGTGGTGTGTTCGCGCCGAGAGCAGCGTCTGCGTGGAGATCTGG 1674
QY 1631 GGAGCCAGGGTGTGGTGTGTTCGCGCCGAGAGCAGCGTCTGCGTGGAGATCTGG 1690
Db 1675 CCAAGTTCTGCACTGCTGATGAGTGTACGTTCGTCGAGCTGCTCAGGTCTTTCTTTT 1734
QY 1691 CCAAGTTCTGCACTGCTGATGAGTGTACGTTCGTCGAGCTGCTCAGGTCTTTCTTTT 1750
Db 1735 ATGTCAGGAGACCACTGCTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGA 1794

QY 1751 ATGTCAGGAGACCACTGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGA 1810
Db 1795 GCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTCGGCGAGCTGT 1854
QY 1811 GCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTCGGCGAGCTGT 1870
Db 1855 CGAAGCAGAGGTCAGGACAGCTCGGAGCAGTCGGAGCAGGCCGCCCTGTCTGACATCCAGACTCC 1914
QY 1871 CGAAGCAGAGGTCAGGACAGCTCGGAGCAGTCGGAGCAGGCCGCCCTGTCTGACATCCAGACTCC 1930
Db 1915 GCTTCATCCCAAGCTGACGCGCTGCGGCCGATGTGAAACATGGACTAGCTCGTGGGAG 1974
QY 1931 GCTTCATCCCAAGCTGACGCGCTGCGGCCGATGTGAAACATGGACTAGCTCGTGGGAG 1990
Db 1975 CCAGAAGCTTCCAGAGAAAGAGGCCGAGCGTCTCACTCGAGGGTGAAGGCACCTGT 2034
QY 1991 CCAGAAGCTTCCAGAGAAAGAGGCCGAGCGTCTCACTCGAGGGTGAAGGCACCTGT 2050
Db 2035 TCAGCGTGTCTAACTAGGAGCGGGCGGCCGCCCTCTCTGGGGCCCTCTGTGTCTGG 2094
QY 2051 TCAGCGTGTCTAACTAGGAGCGGGCGGCCGCCCTCTCTGGGGCCCTCTGTGTCTGG 2110
Db 2095 GCCTGGACGATATCCACAGGGCTTGGCGACCTTCTGTGCTGCTGCTGCGGGCCGAGGACC 2154
QY 2111 GCCTGGACGATATCCACAGGGCTTGGCGACCTTCTGTGCTGCTGCTGCGGGCCGAGGACC 2170
Db 2155 CGCGGCTGAGCTGTACTTGTCAAGTGGATGTGACGGGCGGTAGACACCATCCCC 2214
QY 2171 CGCGGCTGAGCTGTACTTGTCAAGTGGATGTGACGGGCGGTAGACACCATCCCC 2230
Db 2215 AGACAGGCTCAGGAGGTTCATCGCAGCATCATCAAAACCCAGAACACGCTACTGCGTGC 2274
QY 2231 AGACAGGCTCAGGAGGTTCATCGCAGCATCATCAAAACCCAGAACACGCTACTGCGTGC 2290
Db 2275 GTCGGTATGCGGTGGTCCAGAAAGGCCGCCCATGGGACGTCGCCAGGGCCCTTCAAGAGCC 2334
QY 2291 GTCGGTATGCGGTGGTCCAGAAAGGCCGCCCATGGGACGTCGCCAGGGCCCTTCAAGAGCC 2350
Db 2335 ACGTCTCTACTTTGACAGACTCCAGCGTACATGCGACAGTTCTGCTGCTGCTGCGAGG 2394
QY 2351 ACGTCTCTACTTTGACAGACTCCAGCGGTACATGCGACAGTTCTGCTGCTGCTGCGAGG 2410
Db 2395 ANAAGCCCGCTGAGGGATGCGGTGCTATCGAGCAGAGCTCTCTCCCTGAATGAGGCCA 2454
QY 2411 AGACCAAGCCCGCTGAGGGATGCGGTGCTATCGAGCAGAGCTCTCTCCCTGAATGAGGCCA 2470
Db 2455 GCAGTGGCCTTCTGAGCTTCTCTAGCTTCTATGTGCCACACGCGCTGCGCATCAGGG 2514
QY 2471 GCAGTGGCCTTCTGAGCTTCTCTAGCTTCTATGTGCCACACGCGCTGCGCATCAGGG 2530
Db 2515 GCAAGTCTCTACGTCCAGTGCAGGGATCCCGAGGGCTCCATCTCTCCACGCTGCTCT 2574
QY 2531 GCAAGTCTCTACGTCCAGTGCAGGGATCCCGAGGGCTCCATCTCTCCACGCTGCTCT 2590
Db 2575 GCAGCTGTGCTACGCGGACATGGAGAACAGCTGTTTGGGGGATTCGCGGGACGCGG 2634
QY 2591 GCAGCTGTGCTACGCGGACATGGAGAACAGCTGTTTGGGGGATTCGCGGGACGCGG 2650
Db 2635 TGCTCTGCGCTTGGTGGATGATTTCTTGTGTTGACACCTCACCCTACCCACCGGAAAA 2694
QY 2651 TGCTCTGCGCTTGGTGGATGATTTCTTGTGTTGACACCTCACCCTACCCACCGGAAAA 2710
Db 2695 CATTCTCAGGACCTCGTCCGAGGTGTCTCTGAGTATGCTGCTGCTGCTGAACCTTCGGA 2754
QY 2711 CATTCTCAGGACCTCGTCCGAGGTGTCTCTGAGTATGCTGCTGCTGCTGAACCTTCGGA 2770
Db 2755 AGACAGTGGTGAACCTTCCCTGTAGAAGACAGGCGCTTGGTGGCAGCGCTTTTGTTCAGA 2814
QY 2771 AGACAGTGGTGAACCTTCCCTGTAGAAGACAGGCGCTTGGTGGCAGCGCTTTTGTTCAGA 2830
Db 2815 TGCGGCGCCACGCGCTAT 2832
|||||

QY 2831 TGCCGGCCACGGCCTAT 2848
RESULT 10
ID US-08-911-312-18 STANDARD; DNA; UNC; 3855 BP.
AC xxxxxx
DT
DE Sequence 18, Application US/08911312
CC Sequence 18, Application US/08911312
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase
CC NUMBER OF SEQUENCES: 170
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/911.312
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Einhorn, Gregory P.
CC REGISTRATION NUMBER: 38,440
CC REFERENCE/DOCKET NUMBER: 015389-002500US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 3855 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 56...2476

CC OTHER INFORMATION: /product= "hrrr"
CC OTHER INFORMATION: /note= "clone #712562"
SQ SEQUENCE 3855 BP; 651 A; 1300 C; 1226 G; 678 T; 0 OTHER.
Query Match 85.8%; Score 2444; DB 22; Length 3855;
Best Local Similarity 93.6%; Pred. No. 0.00e+00;
Matches 2655; Conservative 0; Mismatches 0; Indels 182; Gaps 1;
Db 1 GCAGCGCTGCTGCTGCTGCGGAGCCCTGGGAGGAGCCCTGGCCCCCGGCCACACCCCGCGATGCC 60
QY 12 GCAGCGCTGCTGCTGCTGCGGAGCCCTGGGAGGAGCCCTGGCCCCCGGCCACACCCCGCGATGCC 71
Db 61 GCGGCTCCCGCTGCGGAGCGCTGCTGCTGCTGCGGAGCCACTACCGGAGGTCT 120
QY 72 GCGGCTCCCGCTGCGGAGCGCTGCTGCTGCTGCGGAGCCACTACCGGAGGTCT 131
Db 121 GCGGCTCCCGCTGCGGAGCGCTGCTGCTGCTGCGGAGCCACTACCGGAGGTCT 180
QY 132 GCGGCTCCCGCTGCGGAGCGCTGCTGCTGCTGCGGAGCCACTACCGGAGGTCT 191
Db 181 GGACCCCGCGGCTTTCGCGCGCTGCTGCTGCTGCGGAGCCACTACCGGAGGTCT 240
QY 192 GGACCCCGCGGCTTTCGCGCGCTGCTGCTGCTGCGGAGCCACTACCGGAGGTCT 251
Db 241 ACGGCCCGCGGCTTTCGCGCGCTGCTGCTGCTGCGGAGCCACTACCGGAGGTCT 300
QY 252 ACGGCCCGCGGCTTTCGCGCGCTGCTGCTGCTGCGGAGCCACTACCGGAGGTCT 311
Db 301 CCAGGTGCTGCGGAGCGCTGCTGCTGCTGCGGAGCCACTACCGGAGGTCT 360
QY 312 CCAGGTGCTGCGGAGCGCTGCTGCTGCTGCGGAGCCACTACCGGAGGTCT 371
Db 361 GCTGCTGCGGAGCGCGGCTGCTGCTGCTGCGGAGCCACTACCGGAGGTCT 420
QY 372 GCTGCTGCGGAGCGCGGCTGCTGCTGCTGCGGAGCCACTACCGGAGGTCT 431
Db 421 CTGCCCCAACACGGTGACGACGCTGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 480
QY 432 CTGCCCCAACACGGTGACGACGCTGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 491
Db 481 CCGGCTGCGGAGCGCGCTGCTGCTGCTGCGGAGCGGCTGCTGCTGCTGCT 540
QY 492 CCGGCTGCGGAGCGCGCTGCTGCTGCTGCGGAGCGGCTGCTGCTGCTGCT 551
Db 541 GGCTCCCGAGCTGCGGCTGCTGCTGCTGCGGAGCGGCTGCTGCTGCTGCT 600
QY 552 GGCTCCCGAGCTGCGGCTGCTGCTGCTGCGGAGCGGCTGCTGCTGCTGCT 611
Db 601 TCAGGCGCGGCGGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 660
QY 612 TCAGGCGCGGCGGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 671
Db 661 CTGGAACCATAGGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 720
QY 672 CTGGAACCATAGGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 731
Db 721 GAGGCGCGGCGGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 780
QY 732 GAGGCGCGGCGGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 791
Db 781 TGCCCTGAGCGGCGGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 840
QY 792 TGCCCTGAGCGGCGGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 851
Db 841 GCGTGGACCGGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 900
QY 852 GCGTGGACCGGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 911
Db 901 CACCTCTTTGGAGGCTGCGCTGCTGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCT 960
QY 912 CACCTCTTTGGAGGCTGCGCTGCTGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCT 971
Db 961 GCACACGCGGCGGCTGCGGAGCGGCTGCTGCTGCGGAGCGGCTGCTGCTGCT 1020

QY 972 GACACAGGGGGCCCCCATCACATCGGGCCACACAGTCCCTCGGACACGCCCTGTGCC 1031
DB 1021 CCCGGTGACGCCGAGACCAAGCACTTCTCTACTCTCAGCGCAACAAGGACAGCTCGG 1080
QY 1032 CCCGGTGACGCCGAGACCAAGCACTTCTCTACTCTCAGCGCAACAAGGACAGCTCGG 1091
DB 1081 GCCCTCCTTCTCTACTCAGCTCTCTAGGCCACAGCTGACTGGCGCTCGGAGGCTGGTGA 1140
QY 1092 GCCCTCCTTCTCTACTCAGCTCTCTAGGCCACAGCTGACTGGCGCTCGGAGGCTCGTGA 1151
DB 1141 GACCATCTTCTGGGTTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCGCT 1200
QY 1152 GACCATCTTCTGGGTTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCGCT 1211
DB 1201 GCCCAGCGCTACTGGCAAAATCGGCCCTGTTTCTGAGCTGCTTGGGAACACAGCGCA 1260
QY 1212 GCCCAGCGCTACTGGCAAAATCGGCCCTGTTTCTGAGCTGCTTGGGAACACAGCGCA 1271
DB 1261 GTGCCCTTACGGGGTGCTCTCAAGAGCACTGCCCGCTGCGAGCTGGGTACACCCAGC 1320
QY 1272 GTGCCCTTACGGGGTGCTCTCAAGAGCACTGCCCGCTGCGAGCTGGGTACACCCAGC 1331
DB 1321 AGCCGGTGCTGTGCCGGGAGAACCCAGGGCTCTGTGGCGGCCCGGAGGAGGA 1380
QY 1332 AGCCGGTGCTGTGCCGGGAGAACCCAGGGCTCTGTGGCGGCCCGGAGGAGGA 1391
DB 1381 CACAGACCCCCCTGCTGCTGAGTGTCTCGCCAGCAGCACAGCCCCCTGGCAGGTGA 1440
QY 1392 CACAGACCCCCCTGCTGCTGAGTGTCTCGCCAGCAGCACAGCCCCCTGGCAGGTGA 1451
DB 1441 CGGGTTCTGCGGGCTGCTGCGCGGGCTGTGTGCCCGCCAGGCCCTTGGGGTCCAGGCA 1500
QY 1452 CGGGTTCTGCGGGCTGCTGCGCGGGCTGTGTGCCCGCCAGGCCCTTGGGGTCCAGGCA 1511
DB 1501 CAACGAAGCCCTTCTCAGAACACCAAGAGTCTACTCCCTGGGGAACATGCCAA 1560
QY 1512 CAACGAAGCCCTTCTCAGAACACCAAGAGTCTACTCCCTGGGGAACATGCCAA 1571
DB 1561 GCTCTCGTGCAGGAGCTGACGTGGAAGATGACGCTGGGGACTGCGCTTGGCTGGGAG 1620
QY 1572 GCTCTCGTGCAGGAGCTGACGTGGAAGATGACGCTGGGGACTGCGCTTGGCTGGGAG 1631
DB 1621 GAGCCAGGGGTGGGTGTTCCGGCGCAGAGCACGCTCTGCGTGAGGAGATCCTGCG 1680
QY 1632 GAGCCAGGGGTGGGTGTTCCGGCGCAGAGCACGCTCTGCGTGAGGAGATCCTGCG 1691
DB 1681 CAAGTCTCTGCACTGGCTGATGAGTGTGACGTCTCGAGCTGCTCAGGTCTTCTTTA 1740
QY 1692 CAAGTCTCTGCACTGGCTGATGAGTGTGACGTCTCGAGCTGCTCAGGTCTTCTTTA 1751
DB 1741 TGTACGGAGACACGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAG 1800
QY 1752 TGTACGGAGACACGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAG 1811
DB 1801 CAAGTTGAAAGCATTTGGAATCAGACAGCACTTGAAGAGGTTGACGCTCGGGAGCTGTC 1860
QY 1812 CAAGTTGAAAGCATTTGGAATCAGACAGCACTTGAAGAGGTTGACGCTCGGGAGCTGTC 1871
DB 1861 GGAAGCAGAGGTTCAGGCACATCGGGAACCCAGGCCCTGCTGACGCTCCAGACTCGG 1920
QY 1872 GGAAGCAGAGGTTCAGGCACATCGGGAACCCAGGCCCTGCTGACGCTCCAGACTCGG 1931
DB 1921 CTTTCATCCCAAGCCTGACGGCTGGCGCGGATTTGTAACATGGACTAGCTCGTGGGAGC 1980
QY 1932 CTTTCATCCCAAGCCTGACGGCTGGCGCGGATTTGTAACATGGACTAGCTCGTGGGAGC 1991
DB 1981 CAGAAGCTTCCGAGAGAAAGAGGCCGAGGCTCTCACCCTCAGGGGTGAAGGCACTGTT 2040
QY 1992 CAGAAGCTTCCGAGAGAAAGAGGCCGAGGCTCTCACCCTCAGGGGTGAAGGCACTGTT 2051
DB 2041 CAGCGTCTCACTAGGAGCGGGCGGGCGGCCCTCTCTGGGCGCTCTGTGCTGGG 2100

QY 2052 CAGCGTCTCACTACGAGCGGGCGGGCCCGGCGCTCTCTGGCGGCTCTGTGCTGGG 2111
DB 2101 CCTGGACGATATCACAGGGCTGGCGACCTTCTGCTGCGTGTGCGGGCCCGAGGACCC 2160
QY 2112 CCTGGACGATATCACAGGGCTGGCGACCTTCTGCTGCGTGTGCGGGCCCGAGGACCC 2171
DB 2161 GCCCGCTCAGCTGACTTGTCAAGGTGGATGTGACGGCGCGGTACGACACCATCCGCCA 2220
QY 2172 GCCCGCTCAGCTGACTTGTCAAGGTGGATGTGACGGCGCGGTACGACACCATCCGCCA 2231
DB 2221 GGACAGGCTCAGGAGTCTCGCAGCATCATCAAAACCCAGAACACGTACTGCGTGGC 2280
QY 2232 GGACAGGCTCAGGAGTCTCGCAGCATCATCAAAACCCAGAACACGTACTGCGTGGC 2291
DB 2281 TCGGTATGCCGTGTCAGAGGCCGCCCATGGGCACTGCGCAAGGCCCTTCAAGAGCCA 2340
QY 2292 TCGGTATGCCGTGTCAGAGGCCGCCCATGGGCACTGCGCAAGGCCCTTCAAGAGCCA 2351
DB 2341 CGTC----- 2344
QY 2352 CGTCTCTACCTTGACAGACTCCAGCCGTACATGCGACAGTTGCTGGCTCACGAGGA 2411
DB 2344 ----- 2344
QY 2412 GACACGCCGCTGAGGGATGCCGTCGTCATCGAGCAGAGCTCTCCCTGAAATGAGGCCAG 2471
DB 2344 ----- 2344
QY 2472 CAGTGGCCTCTTCGAGCTTCTTACGCTTCTATGTGCCACACGCCGTCGCGCATCAGGGG 2531
DB 2345 -----CTACGCTCAGTCCAGGGGATCCCGAGGGTCCATCTCTCCAGCTGCTCTG 2398
QY 2532 CAACTCTCTACGTCAGTCCAGGGGATCCCGAGGGTCCATCTCTCCAGCTGCTCTG 2591
DB 2399 CAGCCTGCTACGCGCATGGAACAAGCTGTTTCGGGGATTCGCGGGAGCGGCT 2458
QY 2592 CAGCCTGCTACGCGCATGGAACAAGCTGTTTCGGGGATTCGCGGGAGCGGCT 2651
DB 2459 GCTCTGCGTTGTTGATGATTTCTTGTGTGACACCTCACCTCACCCACGCCGAAAC 2518
QY 2652 GCTCTGCGTTGTTGATGATTTCTTGTGTGACACCTCACCTCACCCAGCGGAAAC 2711
DB 2519 CTTCTCAGGACCTGTCGAGGTGTCCTGAGTATGGCTGCGTGGTGAACCTTCGGGAA 2578
QY 2712 CTTCTCAGGACCTGTCGAGGTGTCCTGAGTATGGCTGCGTGGTGAACCTTCGGGAA 2771
DB 2579 GACAGTGTGAACCTTCCCTGTAGAGACGAGGCCCTGGTGGCAGCGGCTTTGTTGAT 2638
QY 2772 GACAGTGTGAACCTTCCCTGTAGAGACGAGGCCCTGGTGGCAGCGGCTTTGTTGAT 2831
DB 2639 GCCGGCCACCGCCTAT 2655
QY 2832 GCCGGCCACCGCCTAT 2848
RESULT 11
ID US-08-912-951-4 STANDARD; DNA; UNC; 3855 BP.
AC xxxxxx
DT
DE Sequence 4, Application US/08912951
CC Sequence 4, Application US/08912951
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC TITLE OF INVENTION: THERAPEUTIC METHODS
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:

Db 1201 GCCCCAGGCTACTGGCAAAATCGGCCCCCTGTTTCTGGAGCTGCTTGGGAACACAGCGCA 1260
QY 1212 GCCCCAGGCTACTGGCAAAATCGGCCCCCTGTTTCTGGAGCTGCTTGGGAACACAGCGCA 1271
Db 1261 GTGCCCCAGGGGTGCTCCTCAAGAGCGACTGCCGCTGCGAGCTCGGTTCACCCAGC 1320
QY 1272 GTGCCCCAGGGGTGCTCCTCAAGAGCGACTGCCGCTGCGAGCTCGGTTCACCCAGC 1331
Db 1321 AGCCGGTGTGTGTCGCCGGAGAGCCCCAGGGCTCTGTGGCGGCCCCCGAGGAGGA 1380
QY 1332 AGCCGGTGTGTGTCGCCGGAGAGCCCCAGGGCTCTGTGGCGGCCCCCGAGGAGGA 1391
Db 1381 CACAGACCCCGCTGCTGTCAGCTGCTCGCCACAGACAGCCCTGGCAGTGT 1440
QY 1392 CACAGACCCCGCTGCTGTCAGCTGCTCGCCACAGACAGCCCTGGCAGTGT 1451
Db 1441 CGGCTTGTGCGGGCTGCTGCGCGGCTGGTGCCGCCAGGCTCTGGGGCTCCAGGCA 1500
QY 1452 CGGCTTGTGCGGGCTGCTGCGCGGCTGGTGCCGCCAGGCTCTGGGGCTCCAGGCA 1511
Db 1501 CAAGAACGCGCTTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGATGCCAA 1560
QY 1512 CAAGAACGCGCTTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGATGCCAA 1571
Db 1561 GCTCTCGCTCAGAGCTGAGCTGGAAGATCAGCTGCGGGACTGCGCTTGGCTGCCAG 1620
QY 1572 GCTCTCGCTCAGAGCTGAGCTGGAAGATCAGCTGCGGGACTGCGCTTGGCTGCCAG 1631
Db 1621 GAGCCAGGGGTGGCTGTGTCGGCGGAGAGACCGCTCTGGTGAGGAGATCCCTGGC 1680
QY 1632 GAGCCAGGGGTGGCTGTGTCGGCGGAGAGACCGCTCTGGTGAGGAGATCCCTGGC 1691
Db 1681 CAAGTTCTCTGACCTGCTGATGAGTGTACGTGCTGAGCTGCTCAGGCTTCTTTT 1740
QY 1692 CAAGTTCTCTGACCTGCTGATGAGTGTACGTGCTGAGCTGCTCAGGCTTCTTTT 1751
Db 1741 TGTACGAGACACAGCTTTCAAAGACAGCTCTTTTACCGGGAAGAGTGTCTGGAG 1800
QY 1752 TGTACGAGACACAGCTTTCAAAGACAGCTCTTTTACCGGGAAGAGTGTCTGGAG 1811
Db 1801 CAAGTTCAAAGCATTTGGAATCAGACAGCTTGAAGAGGTGAGCTGCGGGAGCTGTC 1860
QY 1812 CAAGTTCAAAGCATTTGGAATCAGACAGCTTGAAGAGGTGAGCTGCGGGAGCTGTC 1871
Db 1861 GGAAGCAGAGCTCAGGAGCATCGGGAAGCGCCGCTGCTGACGCTCCAGACTCCG 1920
QY 1872 GGAAGCAGAGCTCAGGAGCATCGGGAAGCGCCGCTGCTGACGCTCCAGACTCCG 1931
Db 1921 CTTCATCCCAAGCTCAGCGGCTGCGGCGATTTGTGAACATGAGCTACGTCGTGGAGC 1980
QY 1932 CTTCATCCCAAGCTCAGCGGCTGCGGCGATTTGTGAACATGAGCTACGTCGTGGAGC 1991
Db 1981 CAGAACGTTCCGAGAGAAAGAGGCGCGAGCTCTCACCTCAGAGGTGAAGCAGCTGT 2040
QY 1992 CAGAACGTTCCGAGAGAAAGAGGCGCGAGCTCTCACCTCAGAGGTGAAGCAGCTGT 2051
Db 2041 CAGCGTCTCAACTACAGGCGCGCGCGCCGCTCTCTGGCGCCCTCTGTGCTGGG 2100
QY 2052 CAGCGTCTCAACTACAGGCGCGCGCGCCGCTCTCTGGCGCCCTCTGTGCTGGG 2111
Db 2101 CTTGAGCATATCAGAGGCTGCGGACCTTCGTGCTGCGGTGCGGGCCCAAGGACC 2160
QY 2112 CTTGAGCATATCAGAGGCTGCGGACCTTCGTGCTGCGGTGCGGGCCCAAGGACC 2171
Db 2161 GCCGCTGAGCTGACTTGTCAAGGTGGATGTGACGGGCGCTACGACACCATCCCCCA 2220
QY 2172 GCCGCTGAGCTGACTTGTCAAGGTGGATGTGACGGGCGCTACGACACCATCCCCCA 2231
Db 2221 GGACAGGCTCAGGAGGTCTATCGGAGCATCATCAAAACCCAGAACAGTACTCGTGGC 2280
QY 2232 GGACAGGCTCAGGAGGTCTATCGGAGCATCATCAAAACCCAGAACAGTACTCGTGGC 2291
Db 2281 TCGGTATGCCGTGTCTCCAGAGGCGGCCCATGGGACAGCTCCGCAAGGCTTCAAGAGCCA 2340

QY 2292 TCGGTATGCCGTGGTCCAGAAAGCCGCCCATGGCGACGTCGCCAAGGCCTTCAAGAGCCA 2351
Db 2341 CGTC----- 2344
QY 2352 CGTCTCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTGCTGGCTCACCTGCAGGA 2411
Db 2344 ----- 2344
QY 2412 GACCAGCCGCTGAGGGATGCCGTCGTCATCGAGCAGAGCTCCTCCTGAATGAGGCCAG 2471
Db 2344 ----- 2344
QY 2472 CAGTGGCTCTTCGAGCGTCTTCCTACGTTTCATGTGCGACACGCGCTGGGCATCAGGGG 2531
Db 2345 -----CTAGCTCAGTGGCAGGGGATCCGCGAGGCTCCATCTCTCCACGCTGCTCTG 2398
QY 2532 CAAGTCTCAGTCCAGTCCGCGAGGGATCCGCGAGGCTCCATCTCTCCACGCTGCTCTG 2591
Db 2399 CAGCCTGTGTACGGGACATGGAGAACAGCTGTTTGGGGGATTCGGGGGACGGGCT 2458
QY 2592 CAGCCTGTGTACGGGACATGGAGAACAGCTGTTTGGGGGATTCGGGGGACGGGCT 2651
Db 2459 GCTCCTGCGTTCGATGATTTCTTGTGTGACACCTCACCCTCACCCACCGGAAAC 2518
QY 2652 GCTCCTGCGTTCGATGATTTCTTGTGTGACACCTCACCCTCACCCACCGGAAAC 2711
Db 2519 CTTCTCAGGACCCCTGCTCGAGGTGCTCCGATGATGGCTGGTGGTGAACCTGGCGAA 2578
QY 2712 CTTCTCAGGACCCCTGCTCGAGGTGCTCCGATGATGGCTGGTGGTGAACCTGGCGAA 2771
Db 2579 GACAGTGGTGAACCTCCCTGTAGAAGCAGGCGCCTGGGTGGCAGCGCTTTTGTTCAGAT 2638
QY 2772 GACAGTGGTGAACCTCCCTGTAGAAGCAGGCGCCTGGGTGGCAGCGCTTTTGTTCAGAT 2831
Db 2639 GCCGGCCACGGCTAT 2655
QY 2832 GCCGGCCACGGCTAT 2848

RESULT 12
ID US-08-912-951-6 STANDARD; DNA; UNC; 4200 BP.
AC xxxxxx
DT
Sequence 6, Application US/08912951
CC Sequence 6, Application US/08912951
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435

QY 1467 CTGCTGCGCGGCTGTGCCCCAGGCGCTCTGGGGCTCCAGGCACACAGAACGCCGCTT 1526
Db 3877 CCTAGAACACCAAGATTCATCTCCCTGGGGAAGCATGCCAAGCTCTCGTGCAGGA 3936
QY 1527 CCTAGGAACACCAAGATTCATCTCCCTGGGGAAGCATGCCAAGCTCTCGTGCAGGA 1586
Db 3937 GCTGACGTGGAGATGACGTCGCGGACTGCGCTTGGCTGCGCAGGAGCCAGG 3990
QY 1587 GCTGACGTGGAGATGACGTCGCGGACTGCGCTTGGCTGCGCAGGAGCCAGG 1640

RESULT 13
ID US-09-042-460-1 STANDARD; DNA; UNC; 3496 BP.
AC xxxxxx

DE Sequence 1, Application US/09042460
CC Sequence 1, Application US/09042460
CC GENERAL INFORMATION:
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Allsopp, Richard
CC APPLICANT: Depinho, Ronald
CC APPLICANT: Greenberg, Roger
CC TITLE OF INVENTION: Mouse Telomerase Reverse Transcriptase
CC NUMBER OF SEQUENCES: 101
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/042,460
CC FILING DATE: 16-MAR-1998
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/974,549
CC FILING DATE: 19-NOV-1997

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/974,584
CC FILING DATE: 19-NOV-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/979,742
CC FILING DATE: 26-NOV-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Einhorn, Gregory P.
CC REGISTRATION NUMBER: 38,440
CC REFERENCE/DOCKET NUMBER: 015389-003110US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 3496 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: -
CC LOCATION: 1..3496
CC OTHER INFORMATION: /note= "mouse telomerase reverse
CC OTHER INFORMATION: transcriptase (mTRT) CDNA clone"
CC FEATURE:
CC NAME/KEY: misc_feature
CC LOCATION: 10..3435
CC OTHER INFORMATION: /note= "mouse telomerase reverse
CC OTHER INFORMATION: transcriptase (mTRT) CDNA"
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 39..3404
CC OTHER INFORMATION: /product= "mouse telomerase reverse
CC OTHER INFORMATION: transcriptase (mTRT)"
CC SEQUENCE 3496 BP; 762 A; 998 C; 928 G; 808 T; 0 OTHER.

Query Match 29.0%; Score 825; DB 23; Length 3496;
Best Local Similarity 67.1%; Pred. No. 0.00e+00;
Matches 1901; Conservative 0; Mismatches 838; Indels 93; Gaps 5;

Db 23 CCGCGCTTGAGACATGACCGCGCTCTCTGTTGCCCGCGGTGCGCTCTCTGCTGCG 82
QY 51 CCGCGCACACCCCGCGATGCGCGCTGCCCGCTCCCGAGCGCTGCGCTCCCTGCTCG 110
Db 83 CAGCCGATACCGGAGGTGCGCGCTGGCAACCTTTGTGCGGCGCTTGGGCGCGAGGG 142
QY 111 CAGCACATACCGGAGGTGCTGCGCTGGCCACCGTTGTCGCGCGCTTGGGCGCGCGG 170
Db 143 CAGCGGCTTGTGCAACCCCGGAGACTCTACCGACATTTGTTGTTGCCCAATGCT 202
QY 171 CTGCGGCTGTGTCAGCGCGGAGCCCGCGGCTTCCGCGCGCTGTTGGCCAGTGCCT 230
Db 203 AGTGTGATGACCTGGGGCTCAGAGCTTCCACCTTCCGCGACCTTTCTTCCACAGGTGC 262
QY 231 GGTGTGCTGCTGGGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 290
Db 263 ATCCCTGAAGAGCTGTTGGCGGAGGTGTCAGAGACTCTGCGAGCGCAGCAGAGAGAA 322
QY 291 CTGCTGAAGAGCTGTTGGCGGAGGTGTCAGAGCGCTGTCAGAGCGCGCGCGCGGAGAA 350
Db 323 CGTGTGCTTTTGGCTTTGAGCTGCTTAACAGAGCGCAGAGCGCGCGCGCTTCCATGCGCTT 382
QY 351 CGTGTGCTTTCGCTTCGCGCTGCTGGACGCGGCGCGCGCGCGCGCGCGCGCGCGCTT 410
Db 383 CACTAGTAGCGTGGCTAGCTACTTGTCCCAACACACTGTTATTGAGACCTTGCCTGTCAGTGG 442
QY 411 CACCACAGCGTGGCGAGCTACCTGCTCCCAACACACAGCTGACCGACGCACTTGCAGGAGCGG 470
Db 443 TGCATGGATGCTACTGTTGAGCGGAGTGGGCGAGCACTGCTGCTTACCTGCTGGGACA 502
QY 471 GCGGTGGGGGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 530

QY 2655 CTTGGTTGGTGATTTCTTCTGGTGACACCTCACCTCACCCACGCGAARACCTT 2714
Db 2666 CTTAGCACCTGGTCCATGCGGTTCTCGAGATGCGGTGATGATAAATCTGAGAGAC 2725
QY 2715 CCTAGGACCTGGTCCGAGGTGCTCCCTGAGTATGCTGCGTGGTGAACCTTGGCGAAGAC 2774
Db 2726 AGTGGTGAACCTCCCTGGAGCCCTGTACCTGGTGGTGCAGCTCCATACCACTGCC 2785
QY 2775 AGTGGTGAACCTCCCTGTAGAAGACGAGGCCCTGGGTGGCAGCGCTTTTGTTCAGATGCC 2834
Db 2786 TGCTCACTGCCCT 2797
QY 2835 GGCCACGCGCT 2846

RESULT 14

ID US-09-026-981-36 STANDARD; DNA; UNC; 3346 BP.
AC xxxxxx
DT
DE Sequence 36, Application US/09026981
CC Sequence 36, Application US/09026981
CC GENERAL INFORMATION:
CC APPLICANT: Counter, Christopher M.
CC APPLICANT: Meyerson, Matthew
CC APPLICANT: Weinberg, Robert A.
CC TITLE OF INVENTION: Telomerase Catalytic Subunit Gene and
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Militia Drive
CC CITY: Lexington
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/026,981
CC FILING DATE: 20-FEB-1998
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/064,322
CC FILING DATE: 30-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/055,762
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/054,549
CC FILING DATE: 01-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/047,151
CC FILING DATE: 20-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/038,750
CC FILING DATE: 20-FEB-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WHI97-11p4AM
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 781-861-6240
CC TELEFAX: 781-861-9540
CC INFORMATION FOR SEQ ID NO: 36:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 3346 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 3346 BP; 617 A; 1027 C; 965 G; 735 T; 2 OTHER.

Query Match 28.7%; Score 816; DB 23; Length 3346;
Best Local Similarity 84.9%; Pred. No. 0.00e+00;
Matches 1027; Conservative 0; Mismatches 0; Indels 182; Gaps 1;
Db 971 GGGTTGGCTGTGTTCCGGCCCGCAGACAGCACCTCTGCGTGAGAGATCCTGCGCCAAAGTTCC 1030
QY 1640 GGGTTGGCTGTGTTCCGGCCCGCAGACAGCACCTCTGCGTGAGAGATCCTGCGCCAAAGTTCC 1699
Db 1031 TGCACCTGGCTGATGAGTGTGTACGTGCTGAGCTGCTCAGGTCTTCTTTTATGTACAGG 1090
QY 1700 TGCACCTGGCTGATGAGTGTGTACGTGCTGAGCTGCTCAGGTCTTCTTTTATGTACAGG 1759
Db 1091 AGACCACGTTTCAAAAGAACAGAGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGC 1150
QY 1760 AGACCACGTTTCAAAAGAACAGAGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGC 1819
Db 1151 AAAGCATTTGGAATCAGACAGCACCTTGAAGAGGTGACGTGCGGAGAGTGTCTGGAAGCAG 1210
QY 1820 AAAGCATTTGGAATCAGACAGCACCTTGAAGAGGTGACGTGCGGAGAGTGTCTGGAAGCAG 1879
Db 1211 AGGTACGACGATCGGGAAGCCAGGCCCTCTGTGACGTCCAGACTCCGCTTTCATCC 1270
QY 1880 AGGTACGACGATCGGGAAGCCAGGCCCTCTGTGACGTCCAGACTCCGCTTTCATCC 1939
Db 1271 CCAAGCCTGACGGGCTGCGGCCGATTTGTAACATGGACTAGCTGCTGGAGGCCAGAACGT 1330
QY 1940 CCAAGCCTGACGGGCTGCGGCCGATTTGTAACATGGACTAGCTGCTGGAGGCCAGAACGT 1999
Db 1331 TCCGAGAGAAAAGAGGCCCGAGGCTCTACCTCGAGGGTGAAGCAGCTGTTTACGCGTGC 1390
QY 2000 TCCGAGAGAAAAGAGGCCCGAGGCTCTACCTCGAGGGTGAAGCAGCTGTTTACGCGTGC 2059
Db 1391 TCAACTAGACGCGGGCGCGGCCCTCTGCGGCCCTCTGTGCTGGGCTTGGACG 1450
QY 2060 TCAACTAGACGCGGGCGCGGCCCTCTGCGGCCCTCTGTGCTGGGCTTGGACG 2119
Db 1451 ATATCCACAGGGCTGCGGCACCTTCTGTGCTGCGGTGCGGGGCCAGAACCCGCGCCTG 1510
QY 2120 ATATCCACAGGGCTGCGGCACCTTCTGTGCTGCGGTGCGGGGCCAGAACCCGCGCCTG 2179
Db 1511 AGCTGTACTTTGTCAAGGTGGATGTGACGGGCGGTAGACACCATCCCGCAGACAGGC 1570
QY 2180 AGCTGTACTTTGTCAAGGTGGATGTGACGGGCGGTAGACACCATCCCGCAGACAGGC 2239
Db 1571 TCACGGAGGTCTATCGCCAGCATCATCAAAACCCAGAACACACTACTGCTGGTGGGTATG 1630
QY 2240 TCACGGAGGTCTATCGCCAGCATCATCAAAACCCAGAACACACTACTGCTGGTGGGTATG 2299
Db 1631 CCGTGGTCCAGAGGCCCGCCCATGGGACGTCGCGAAGCCCTTCAAGAGCCACGTC---- 1686
QY 2300 CCGTGGTCCAGAGGCCCGCCCATGGGACGTCGCGAAGCCCTTCAAGAGCCACGTCCTCTA 2359
Db 1686 ----- 1686
QY 2360 CTTGACAGACCTTCAGCGCTACATGCGACAGTTCTGTGGCTACCTGACGAGACAGGCC 2419
Db 1686 ----- 1686
QY 2420 CGCTAGGGATGCCGCTCTCATCGACAGAGCTCTCTCCTGAATGAGGCCAGCAGTGCC 2479
Db 1687 -----CT 1688
QY 2480 TCTTCGAGCTCTTCTACGCTTTCATGTGCCACACGCGCGTGGCATCAGGGGCAAGTCT 2539
Db 1689 AGTCCAGTGCAGGGGATCCCGCAGGCTCTCATCTCTCCACGCTGCTCTGACGCTGT 1748
QY 2540 AGTCCAGTGCAGGGGATCCCGCAGGCTCTCATCTCTCCACGCTGCTCTGACGCTGT 2599
Db 1749 GCTACGGGACATGGAGAACAAAGCTGTTTGGGGGATTCGCGGAGCGGCTGCTCTGTC 1808
QY 2600 GCTACGGGACATGGAGAACAAAGCTGTTTGGGGGATTCGCGGAGCGGCTGCTCTGTC 2659
Db 1809 GTTTGGTGATGATTTCTTGTGGTGACACCTCACCTCACCCACGCGAAGAACCTTCTCTCA 1868

QY 2469 CAGCAGTGGCCTCTTCACGCTCTTCCTACGCTTCATGTGCCACACGCCGTGCCCATCAG 2528
Db 668 -----CTACGCTCCAGTGGCCAGGGGATCCCCGAGGGCTCCATCCTCTCCACGCTGCT 718
|||||
QY 2529 GGGCAAGTCTCTACGCTCCAGTGGCCAGGGGATCCCCGAGGGCTCCATCCTCTCCACGCTGCT 2588
Db 719 CTGCAGCCTGTGCTACGGCGACATGGAGAACAGCTGTTTGGGGGATTCGGGGGGACGG 778
|||||
QY 2589 CTGCAGCCTGTGCTACGGCGACATGGAGAACAGCTGTTTGGGGGATTCGGGGGGACGG 2648
Db 779 GCTGCTCTCGGTTTGGTGATGATTTCTTGGTGACACCTCACCTCACCCACGCGAA 838
|||||
QY 2649 GCTGCTCTCGGTTTGGTGATGATTTCTTGGTGACACCTCACCTCACCCACGCGAA 2708
Db 839 AACCTTCTCAGGACCTGGTCCGAGGTGTCCTGAGTATGGCTGCGTGGTGAACCTTGC 898
|||||
QY 2709 AACCTTCTCAGGACCTGGTCCGAGGTGTCCTGAGTATGGCTGCGTGGTGAACCTTGC 2768
Db 899 GAAGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGCTTTTGTCA 958
|||||
QY 2769 GAAGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGCTTTTGTCA 2828
Db 959 GATCGCGGCCACGGCCTAT 978
|||||
QY 2829 GATCGCGGCCACGGCCTAT 2848

Search completed: Fri Dec 25 17:47:26 1998
Job time : 9031 secs.

THIS PAGE BLANK (USPTO)

WIPEOUT (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 15:11:23 1998; MasPar time 245.67 Seconds
Tabular output not generated. 540.766 Million cell updates/sec

Title: >US-08-951-733-13
Description: (1-2848) from US08951733.seq (1 of 3)
Perfect Score: 2848
N.A. Sequence: 1 CACGCGTCCGGCAGCGCTG.....GATCGCGGCCGCGCTAT 2848
Comp: GTGCGCAGCGCGTCCGGCAGC.....CTACGCGCGGTCCGGGATA

Scoring table: TABLE default
Gap open 30; Gap extend 1

Mismatch STD : Dbase 0; Query 0

Searched: 88822 seqs, 23323279 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-issued
1:5_COMB 2:PCT9_COMB 3:backfiles1

Statistics: Mean 9.326; Variance 5.654; scale 1.649

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Length | ID | Description | Pred. No. |
|------------|-------------|--------|------|------------------------|-----------|
| 1 | 67 | 2.4 | 7218 | Sequence 14, Applicati | 2.08e-25 |
| 2 | 49 | 1.7 | 7218 | Sequence 14, Applicati | 1.91e-14 |
| 3 | 34 | 1.2 | 215 | Sequence 5, Applicatio | 6.10e-06 |
| 4 | 35 | 1.2 | 215 | Sequence 5, Applicatio | 1.76e-06 |
| 5 | 25 | 0.9 | 66 | Sequence 144, Applicat | 2.32e-01 |
| 6 | 25 | 0.9 | 68 | Sequence 243, Applicat | 2.32e-01 |
| 7 | 25 | 0.9 | 69 | Sequence 142, Applicat | 2.32e-01 |
| 8 | 26 | 0.9 | 74 | Sequence 94, Applicati | 7.68e-02 |
| 9 | 25 | 0.9 | 74 | Sequence 100, Applicat | 2.32e-01 |
| 10 | 26 | 0.9 | 74 | Sequence 100, Applicat | 7.68e-02 |
| 11 | 26 | 0.9 | 74 | Sequence 99, Applicati | 7.68e-02 |
| 12 | 25 | 0.9 | 75 | Sequence 99, Applicati | 2.32e-01 |
| 13 | 25 | 0.9 | 75 | Sequence 99, Applicati | 2.32e-01 |
| 14 | 26 | 0.9 | 81 | Sequence 92, Applicati | 7.68e-02 |
| 15 | 25 | 0.9 | 81 | Sequence 98, Applicati | 2.32e-01 |
| 16 | 26 | 0.9 | 81 | Sequence 98, Applicati | 7.68e-02 |
| 17 | 26 | 0.9 | 81 | Sequence 92, Applicati | 7.68e-02 |
| 18 | 26 | 0.9 | 82 | Sequence 97, Applicati | 7.68e-02 |
| 19 | 26 | 0.9 | 82 | Sequence 97, Applicati | 7.68e-02 |

| | | | | | | | |
|----|----|-----|------|---|------------|------------------------|----------|
| 20 | 25 | 0.9 | 242 | 1 | US-08-273- | Sequence 1, Applicatio | 2.32e-01 |
| c | 21 | 0.9 | 1004 | 2 | PCT-US95-0 | Sequence 7, Applicatio | 2.32e-01 |
| c | 22 | 0.9 | 1386 | 2 | PCT-US95-0 | Sequence 2, Applicatio | 2.32e-01 |
| c | 23 | 0.9 | 1611 | 2 | PCT-US93-0 | Sequence 3, Applicatio | 2.32e-01 |
| c | 24 | 0.8 | 59 | 2 | PCT-US95-1 | Sequence 95, Applicati | 1.97e+00 |
| c | 25 | 0.8 | 65 | 1 | US-08-471- | Sequence 145, Applicat | 6.84e-01 |
| c | 26 | 0.8 | 66 | 2 | PCT-US95-1 | Sequence 93, Applicati | 6.84e-01 |
| c | 27 | 0.8 | 66 | 1 | US-08-471- | Sequence 144, Applicat | 1.97e+00 |
| c | 28 | 0.8 | 68 | 1 | US-08-471- | Sequence 143, Applicat | 6.84e-01 |
| c | 29 | 0.8 | 69 | 1 | PCT-US95-1 | Sequence 142, Applicat | 1.97e+00 |
| c | 30 | 0.8 | 84 | 2 | PCT-US94-0 | Sequence 25, Applicati | 1.97e+00 |
| c | 31 | 0.8 | 84 | 2 | US-08-133- | Sequence 25, Applicati | 1.97e+00 |
| c | 32 | 0.8 | 84 | 1 | US-08-300- | Sequence 120, Applicat | 1.97e+00 |
| c | 33 | 0.8 | 84 | 1 | US-08-299- | Sequence 25, Applicati | 1.97e+00 |
| c | 34 | 0.8 | 225 | 1 | US-08-273- | Sequence 18, Applicati | 1.97e+00 |
| c | 35 | 0.8 | 242 | 1 | US-08-273- | Sequence 1, Applicatio | 6.84e-01 |
| c | 36 | 0.8 | 1404 | 3 | 5480796-8 | Patent No. 5480796 | 1.97e+00 |
| c | 37 | 0.8 | 1404 | 3 | 5171840-8 | Patent No. 5171840 | 1.97e+00 |
| c | 38 | 0.8 | 1640 | 1 | US-07-807- | Sequence 11, Applicati | 1.97e+00 |
| c | 39 | 0.8 | 1640 | 1 | US-08-299- | Sequence 11, Applicati | 1.97e+00 |
| c | 40 | 0.8 | 2061 | 3 | 5171840-1 | Patent No. 5171840 | 1.97e+00 |
| c | 41 | 0.8 | 2061 | 3 | 5480796-1 | Patent No. 5480796 | 1.97e+00 |
| c | 42 | 0.8 | 2253 | 3 | 5457037-2 | Patent No. 5457037 | 1.97e+00 |
| c | 43 | 0.8 | 2454 | 1 | US-07-872- | Sequence 1, Applicatio | 6.84e-01 |
| c | 44 | 0.8 | 2625 | 3 | 5457037-4 | Patent No. 5457037 | 1.97e+00 |
| c | 45 | 0.8 | 3336 | 3 | 5457037-1 | Patent No. 5457037 | 1.97e+00 |

ALIGNMENTS

RESULT 1
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC xxxxxx

Sequence 14, Application US/08232463
Sequence 14, Application US/08232463
Patent No. 5670367

GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:

```
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: PT29pt-F1s
SQ SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

Query Match 2.4%; Score 67; DB 1; Length 7218;
Best Local Similarity 2.4%; Pred. No. 2.08e-25;
Matches 9; Conservative 211; Mismatches 153; Indels 0; Gaps 0;

Db 1055 GGAGCTTCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1114
||||| ||||| : : : : : : : : : : : : : : : : : : : : : : :
Qy 921 GGAGGTCGCTCTGTCAGCGCCACTCCACCACCATCGTGGCGCCGACACGC 980
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1115 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1174
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 981 GGGCCCCCATCCACATCGCGCCACACAGTCCCGGGACAGCGCTTGTCCCGGTGA 1040
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1175 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1234
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1041 CGCCGAGACCAAGCACTCTCTACTCTCAGGCGACAAGGAGCAGCTGCGCCCTCCTT 1100
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1235 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1294
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1101 CCTACTAGCTCTGTCAGGCGCCGCTGACTGCGCTCGGAGGCTCGTGGAGACCATCTT 1160
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1295 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1354
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1161 TCTGGGTTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTGCCCCGCTGCCCGACGG 1220
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1355 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1414
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1221 CTACTGGCAATCGGCCCTCTTCTTGAGCTGCTTGGAAACACGCGCAGTGCCCTTA 1280
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1415 YYYYYYYYYYYY 1427
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1281 CGGGGTGCTCCTC 1293
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC xxxxxx
DT
Sequence 14, Application US/08232463
CC Sequence 14, Application US/08232463
CC Patent No. 5670367
CC GENERAL INFORMATION:
CC APPLICANT: DORNER, F.
CC APPLICANT: SCHEIFLINGER, F.
CC APPLICANT: FALKNER, F. G.
CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22313-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/232.463
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/07/935.313
```

```
CC FILING DATE:
CC APPLICATION NUMBER: EP 91 114 300.6
CC FILING DATE: 26-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703)836-9300
CC TELEX: 899149
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: PT29pt-F1s
SQ SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

Query Match 1.7%; Score 49; DB 1; Length 7218;
Best Local Similarity 2.0%; Pred. No. 1.91e-14;
Matches 6; Conservative 165; Mismatches 122; Indels 0; Gaps 0;

Db 1150 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1209
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 2079 CCGCGCCGCTCGTAGTGTGACAGCGCTGAACAGTGCCTTCACCTCGAGGTGAGACGCTC 2020
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1210 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1269
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 2019 GGGCCTCTTTCTCTCGGAACGTTCTGGTCCACGACGTAGTCCATGTTCAACAATCGG 1960
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1270 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1329
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 1959 CCGCAGCCGCTCAGCTGGGATGAAGCGAGTCTGGAGCTCAGAGGGCGGCTGGC 1900
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1330 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1389
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 1899 TTCCGATGCTGCTGACCTCTGCTCCGACAGCTCCGCGAGCTGCACCCCTCTTCAAGTG 1840
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1390 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1442
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 1839 CTGCTGATTCCAATGCTTGGCAACTGCTCCAGACACTCTTCCGGTAGAAAA 1787
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DT
Sequence 5, Application US/08238163
CC Sequence 5, Application US/08238163
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: STOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEA
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourie and Crew
CC STREET: Steuart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
```

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc_feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 1.2%; Score 34; DB 1; Length 215;
Best Local Similarity 12.9%; Pred. No. 6.10e-06;
Matches 17; Conservative 56; Mismatches 59; Indels 0; Gaps 0;

Db 83 HTNVSADSKVTVDYNSAGTSSNGTGDNRSGADSYGSSKTAAMTSRRTGKTANNAYD 142
QY 130 CTGCGCTGGCCAGCTTCGTCGGCGCTGGGCCCCAGGCTGGCGTGGTGACGCG 189
Db 143 SRNMGASVSGDKNTKKHAKNSADGVSKNNGDRNNRYGTGTGTSNVSNNCGGKNRDKVS 202
QY 190 GGGGACCGCGGGCTTTCCGGCGCTGTGTGCGCCAGTGCCTGGTGGTGGTGGGAC 249
Db 203 SYANKCGGSSC 214
QY 250 GCACGGCGCGCC 261

RESULT 4
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DT Sequence 5, Application US/08238163
DE Sequence 5, Application US/08238163
CC Patent No. 556830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: STOFZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC TITLE OF INVENTION: POLYLACTONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESSES:
CC ADDRESS: Townsend and Townsend Kourie and Crew
CC STREET: Stewart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800

CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc_feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 1.2%; Score 35; DB 1; Length 215;
Best Local Similarity 13.2%; Pred. No. 1.76e-06;
Matches 26; Conservative 76; Mismatches 95; Indels 0; Gaps 0;

Db 6 MSSSVVSRFASCNDRKAKDGTSSWTTDCNRTWGVCDTFTYRYNNDSGHNKYSAN 65
Cp 1615 AGTCCCGACGCTCATCTCCAGCTCAGCTCCTGCAGCAGAGCTTGGCATGCTTCCCA 1556
Db 66 YNTGGNNVGAHYYHTNVSADSKVTVDYNSAGTSSNGTGDNRSGADSYGSSKT 125
Cp 1555 GGGAGATGAACCTTCTGGTGTCTTCTGAGGAAGCGCGCTTCTGTGCTGGAGCCCA 1496
Db 126 AMTSRRTGTANNVDSRNMGDASVSGDRNKKHAKNSADGVSKNNGDRNNRYGTGT 185
Cp 1495 GGCTGGGGCACCAGCGCGCGCAGCAGCGCCGCCACGACCGGTACCTGCGCAGGGC 1436
Db 186 KSNVNNCGGKNRDKVS 202
Cp 1435 TGCTGTGCTGGCGGAGC 1419

RESULT 5
ID US-08-471-052A-144 STANDARD; DNA; UNC; 66 BP.
AC xxxxxx
DT Sequence 144, Application US/08471052A
DE Sequence 144, Application US/08471052A
CC Patent No. 5625033
CC GENERAL INFORMATION:
CC APPLICANT: Kay, B. K.
CC APPLICANT: Fowlkes, D. M.
CC TITLE OF INVENTION: Totally Synthetic Affinity Reagents
CC NUMBER OF SEQUENCES: 166
CC CORRESPONDENCE ADDRESSES:
CC ADDRESS: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/471,052A
CC FILING DATE: 06-JUNE-1995
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872

Query Match 0.9%; Score 26; DB 2; Length 74;
Best Local Similarity 11.4%; Pred. No. 7.68e-02;
Matches 8; Conservative 19; Mismatches 43; Indels 0; Gaps 0;


```

CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
SQ      SEQUENCE 81 BP; 6 A; 6 C; 4 G; 5 T; 60 OTHER.

Query Match      0.9%; Score 25; DB 2; Length 81;
Best Local Similarity 7.7%; Pred. No. 2.32e-01;
Matches 5; Conservative 20; Mismatches 40; Indels 0; Gaps 0;

Db 13 VNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 72
Qy 945 CCACGCCACCCATCGGTGGCGCGCCAGCACCGGGGGCCCCCATCCACATCGGGCC 1004
Db 73 ACCAC 77
Qy 1005 ACCAC 1009

```

Search completed: Fri Dec 25 15:16:36 1998
 Job time : 313 secs.

M P S R E N E H

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 09:43:38 1998; MagPar time 5827.06 Seconds
Tabular output not generated.
726.145 Million cell updates/sec

Title: >US-08-951-733-13
Description: (1-1920) from US08951733.seq (1 of 2)
Perfect Score: 1920
N.A. Sequence: 1 CAGCGCGTCCGGCAGGCGTG.....CCAGGCGCGCGCCCTGCTGACG 1920
Comp: GTCCGCGAGGCGCGTCCGAC.....GTCGCGGCGGCGACGACTGC

Scoring table: TABLE default
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb155
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_lo 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_vl

Database: genbank107
15:gb_ba1 16:gb_ba2 17:gb_htg 18:gb_in 19:gb_om 20:gb_ov
21:gb_pat 22:gb_ph 23:gb_pl1 24:gb_pl2 25:gb_pr1
26:gb_pr2 27:gb_pr3 28:gb_ro 29:gb_st 30:gb_sts 31:gb_sy
32:gb_un 33:gb_vl

Statistics: Mean 11.310; Variance 6.127; scale 1.846

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| SUMMARIES | | | |
|------------|-------|--------------------------|---|
| Result No. | Score | Query Match Length DB ID | Description Pred. No. |
| 1 | 1909 | 99.4 | 4015 27 AF015950 Homo sapiens telomeras 0.00e+00 |
| 2 | 1896 | 98.8 | 8960 31 AF043739 Synthetic construct hu 0.00e+00 |
| 3 | 1895 | 98.7 | 4027 27 AF018167 Homo sapiens telomeras 0.00e+00 |
| 4 | 453 | 23.6 | 3426 28 AF051911 Mus musculus telomeras 0.00e+00 |
| 5 | 451 | 23.5 | 3369 28 AF073311 Mus musculus telomeras 0.00e+00 |
| 6 | 110 | 5.7 | 201 28 AF029235 Mus musculus telomeras 3.62e-55 |
| 7 | 67 | 3.5 | 7218 21 I66494 Sequence 14 from paten 6.05e-25 |
| 8 | 44 | 2.3 | 7218 21 I66494 Sequence 14 from paten 4.31e-10 |
| 9 | 37 | 1.9 | 74371 26 AC005369 Homo sapiens chromosom 5.98e-06 |
| 10 | 34 | 1.8 | 215 21 I28278 Sequence 5 from patent 2.93e-04 |
| 11 | 35 | 1.8 | 215 21 I28278 Sequence 5 from patent 8.13e-05 |
| 12 | 34 | 1.8 | 74371 26 AC005369 Homo sapiens chromosom 2.93e-04 |

| | | | | | | | |
|----|----|-----|-----------|------------|-------------------------|------------------------|----------|
| 13 | 30 | 1.6 | 201 21 | A10158 | S.griseus gene for pre | 4.16e-02 | |
| 14 | 30 | 1.6 | 201 21 | A10161 | Synthetic DNA for prep | 4.16e-02 | |
| c | 15 | 30 | 201 21 | A10162 | Synthetic DNA for prep | 4.16e-02 | |
| c | 16 | 30 | 201 21 | A10159 | S.griseus gene for pre | 4.16e-02 | |
| 17 | 30 | 1.6 | 565 21 | E04076 | gDNA encoding envelope | 4.16e-02 | |
| 18 | 31 | 1.6 | 216021 26 | HUAC004787 | Homo sapiens Chromosom | 1.24e-02 | |
| 19 | 28 | 1.5 | 1663 19 | MVU92534 | Mustela vislon microsat | 4.40e-01 | |
| 20 | 29 | 1.5 | 6011 33 | HS4RSIR22T | Epstein-Barr virus (wi | 1.37e-01 | |
| c | 21 | 28 | 31812 24 | SPUNK4 | S.pombe chromosome I c | 4.40e-01 | |
| c | 22 | 28 | 1.5 | 39073 23 | SPAC2E11 | S.pombe chromosome I c | 4.40e-01 |
| 23 | 29 | 1.5 | 76094 25 | HS179D3A | Human DNA sequence fro | 1.37e-01 | |
| c | 24 | 27 | 1.4 | 565 21 | E04076 | gDNA encoding envelope | 1.38e+00 |
| c | 25 | 27 | 1.4 | 904 33 | HS4BHLF | Epstein-Barr virus Bam | 1.38e+00 |
| 26 | 26 | 1.4 | 1341 20 | CHKDILRB | Gallus domesticus DI-1 | 4.23e+00 | |
| 27 | 26 | 1.4 | 2926 15 | SMAPROBA | S.marcescens proBA ope | 4.23e+00 | |
| 28 | 26 | 1.4 | 2926 15 | SMPROBAG | S.marcescens proBA gen | 4.23e+00 | |
| 29 | 26 | 1.4 | 4767 28 | AF034863 | Rattus norvegicus syna | 4.23e+00 | |
| 30 | 27 | 1.4 | 5050 33 | HS4U2IR2A | Epstein-Barr virus (B9 | 1.38e+00 | |
| 31 | 27 | 1.4 | 5400 15 | SCACTVA | S.coelicolor 6 activa r | 1.38e+00 | |
| 32 | 26 | 1.4 | 12949 16 | AF011922 | Pseudomonas aeruginosa | 4.23e+00 | |
| c | 33 | 27 | 1.4 | 35980 16 | MTY20H10 | Mycobacterium tubercul | 1.38e+00 |
| c | 34 | 26 | 1.4 | 43034 26 | HSN5H6 | Human DNA sequence fro | 4.23e+00 |
| 35 | 27 | 1.4 | 64078 18 | AC004365 | Drosophila melanogaste | 1.38e+00 | |
| c | 36 | 26 | 1.4 | 85132 26 | U82668 | Homo sapiens shox gene | 4.23e+00 |
| 37 | 26 | 1.4 | 123491 17 | HS126A5 | Human DNA sequence *** | 4.23e+00 | |
| 38 | 26 | 1.4 | 125815 17 | HS436C18 | Human DNA sequence *** | 4.23e+00 | |
| 39 | 26 | 1.4 | 138706 17 | AC004946 | *** SEQUENCING IN PROG | 4.23e+00 | |
| 40 | 26 | 1.4 | 147216 27 | AF093117 | Homo sapiens chromosom | 4.23e+00 | |
| 41 | 26 | 1.4 | 147542 17 | HS754E20 | Human DNA sequence *** | 4.23e+00 | |
| 42 | 27 | 1.4 | 172281 33 | EBV | Epstein-Barr virus (EB | 1.38e+00 | |
| 43 | 27 | 1.4 | 184113 33 | HS4B958RAJ | Epstein-Barr virus, ar | 1.38e+00 | |
| 44 | 25 | 1.3 | 69 21 | I41362 | Sequence 142 from pate | 1.26e+01 | |
| 45 | 25 | 1.3 | 201 21 | A10159 | S.griseus gene for pre | 1.26e+01 | |

ALIGNMENTS

| | | | | | | |
|-----------|------------|---|---------|------|-----|-------------|
| 1 | RESULT | AF015950 | 4015 bp | mrna | PRI | 16-AUG-1997 |
| LOCUS | DEFINITION | Homo sapiens telomerase reverse transcriptase (hTERT) mRNA, complete cds. | | | | |
| ACCESSION | NID | AF015950 | | | | |
| KEYWORDS | SOURCE | human. | | | | |
| ORGANISM | REFERENCE | Homo sapiens | | | | |
| | AUTHORS | Nakamura,T.M., Morin,G.B., Chapman,K.B., Weinrich,S.L., Andrews,W.H., Lingner,J., Harley,C.B. and Cech,T.R. | | | | |
| TITLE | JOURNAL | Telomerase catalytic subunit homologs from fission yeast and human | | | | |
| MEDLINE | REFERENCE | Science 277 (5328), 955-959 (1997) | | | | |
| AUTHORS | TITLE | Direct Submission | | | | |
| JOURNAL | FEATURES | Submitted (24-JUL-1997) Geron Corporation, 230 Constitution Drive, Menlo Park, CA 94025, USA | | | | |
| | source | Location/Qualifiers | | | | |
| | | 1. .4015 | | | | |
| | | /organism="Homo sapiens" | | | | |
| | | /db_xref="taxon:9606" | | | | |
| | | /tissue_type="kidney" | | | | |
| | | /map="5p" | | | | |
| | | /dev_stage="embryo" | | | | |
| | | /chromosome="5" | | | | |
| | gene | 1. .4015 | | | | |
| | | /gene="hTERT" | | | | |
| | CDS | 56. .3454 | | | | |
| | | /gene="hTERT" | | | | |
| | | /codon_start=1 | | | | |

|||||
QY 1752 TGTCCAGGACACGCTTTTCAAAGAACAGCTCTTTTCTACCGAAGAGTGCTCGAG 1811
|||||
Db 1801 CAAGTTCGAAGCATTTGAATCAGACAGCACTTGAAGAGGCTGACGTGCGGGAGCTGTC 1860
|||||
QY 1812 CAAGTTCGAAGCATTTGAATCAGACAGCACTTGAAGAGGCTGACGTGCGGGAGCTGTC 1871
|||||
Db 1861 GGAAGCAGAGTCCAGGAGCATTCGGGAAGCAGCCCGCCCTGCTGACG 1909
|||||
QY 1872 GGAAGCAGAGTCCAGGAGCATTCGGGAAGCAGCCCGCCCTGCTGACG 1920
|||||

RESULT 2
LOCUS AF043739 8960 bp DNA SYN 05-APR-1998
DEFINITION Synthetic construct human telomerase catalytic subunit
(hTERT/hEST2) mRNA and 3' Influenza A hemagglutinin tag, complete
sequence.
ACCESSION AF043739
NID g3023054
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
SUBMITTED (20-JAN-1998) Whitehead Institute, 9 Cambridge Center,
Cambridge, MA 02142, USA
FEATURES
source
location/Qualifiers
1. 8960
/organism="synthetic construct"
/db_xref="taxon:32630"
order(1. 1101,4606. 8960)
/note="mammalian expression vector; see GenBank Accession
Number U47120"
/organism="Cloning vector pCI-neo"
/db_xref="taxon:45772"
1102. 4542
/note="telomerase catalytic subunit hTERT/hEST2 mRNA"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5p15.33"
1102. 4542
/gene="hTERT/hEST2"
1147. 4605
/note="telomerase/reverse transcriptase"
/codon_start=1
/transl_table=11
/product="telomerase catalytic subunit"
/db_xref="PID:g3023055"
/translation="MPRAPKRAVRSLRSHYREVLPLATFVRRLGPGWRLVORGDP
AAFLVAQCLVCPWDPAPPPAPSPFQVSLKELVARVLQRLCERGNLAFGEA
LLDGARGPEAFTTSVSYLPTNTDALRGAWGLLLRRVGDVLLVHLLARCALEV
LVAPSCAYQVCPFLYOLGAATQAPPPHAGSPRRRLGRCERAWNHSVREAGVPLGLPA
PGARRGGSASRLPLPKPRRGAAPERTPVCGGSAHPGTRGSDRGFCVWSPA
RPAEATSLGASLGRSHSPVGRQHGAPPTSRPRPDTPCPVYAEKHELFS
SGDEQRPSFLLSSLRPTGARRLVETIFLGRSPWMPGTPRRLPLRPQRYQMRLP
FLELLGHNAQCPYGVLLKTHCPLRAAVTPAAGVAREKPGQSVAAPEEDTPRRLVQ
LLRQSHSPWQYVTRACLRRLVPPGLWGRHNRRLRNKTFISLGHKHAISLQLEL
TKMSVRGCAWLRSPGVCVPAAEHLREELAKFLHLMVSVYVVELLSFFVYET
TFONRLFFYKRSVMSKLSIGIFQHLKRVQLRELSEAEVROHEARPALLTSLRFI
PKPGLRPIYNDYVVGARTREKRAERLTSRKALFSLVNLTERARRPGLLGNVIG
LDDTHRAWRTFVLVRQADPPPELYFVKVDVGTAYDTPQDRLTEVLAISLIKPNITVC
VRYAVYVQKAHGHVHRKAFKSHVSTLTLDLPYMRQFVAHLQETSPLRDAVVIQSSSL

NEASSCLFDVFLRFMCHHAVIRGKSYVOCQIPQSGIISLTLLCSLCYGDMEKLFAG
IRRDGLLLLDVDFLLVTPHLTHAKTFLRLTVRGVPEYGCYVNLRLTIRVNFVPEDEAL
GTGAFVQMPAHGLFPMCGLLDTRTLEVDQSDYSYARTSIRASLTFRNFKGRNNRR
KLFGLVRLKCHSLFLDLQVNSLTQVCTNTYKILLQYRHFACVLAQLPFGHQVKNRP
FLFRVSDTASLICYSLKAKNAGMSLGAAGPLPSEAVQWLCHQALLKLRHRT
VYPLGLSLRTAQTQLSRKLPGLTTLTALEAANPALPSDFKTLILDYVDYDYPDAGYPYD
VPDYAG"
source 4543. 4605
/note="engineered based on the epitope for the
anti-influenza virus hemagglutinin monoclonal antibody
12CA5; similar to part of the Influenza A hemagglutinin
gene sequence in GenBank Accession Number M29257"
/organism="synthetic construct"
/db_xref="taxon:32630
BASE COUNT 1886 a 2551 c 2484 g 2038 t 1 others
ORIGIN
Query Match 98.8%; Score 1896; DB 31; Length 8960;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1898; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 1101 CGTCCTGCTGCGCAGCTGGGAAGCCCTGGCCCGCCGACCCCGCGATGCGCGCTCC 1160
|||||
QY 21 CGTCCTGCTGCGCAGCTGGGAAGCCCTGGCCCGCCGACCCCGCGATGCGCGCTCC 80
|||||
Db 1161 CCGCTGCCGAGCGCTGCGCTCCCTGTGCGAGCCACTACCGGAGGTGCTGCGCTGGC 1220
|||||
QY 81 CCGCTGCCGAGCGCTGCGCTCCCTGTGCGCAGCACTACCGGAGGTGCTGCGCTGGC 140
|||||
Db 1221 CACGTTGCTGGCGCGCTGGGCCCGCAGGGCTGGCGCTGGTGACGCGGGACCGGC 1280
|||||
QY 141 CACGTTGCTGGCGCGCTGGGCCCGCAGGGCTGGCGCTGGTGACGCGGGACCGGC 200
|||||
Db 1281 GGCYTTCCGCGCGCTGGTGCGCCAGTGCCTGCTGCTGGCTGCGCTGGGACGACGCGCGCC 1340
|||||
QY 201 GGCYTTCCGCGCGCTGGTGCGCCAGTGCCTGCTGCTGGCTGGGACGACGCGCGCC 260
|||||
Db 1341 CCCCGCGCGCGCTTCCTTCGCGCCAGGTGCTGCTGCTGAAGAGAGTGTGTGGCCCCAGTGT 1400
|||||
QY 261 CCCCGCGCGCGCTTCCTTCGCGCCAGGTGCTGCTGCTGAAGAGAGTGTGTGGCCCCAGTGT 320
|||||
Db 1401 GCAGAGCTGTGCGAGCGCGCGCGGAAGAGCTGCTGGCCTTCGCGCTTCGCGCTGCTGGA 1460
|||||
QY 321 GCAGAGCTGTGCGAGCGCGCGCGGAAGAGCTGCTGGCCTTCGCGCTTCGCGCTGCTGGA 380
|||||
Db 1461 CGGGCGCGCGCGCGCGCGCGCGAGGCTTCACCCACAGCGCTGCGCAGCTACCTGCCCAA 1520
|||||
QY 381 CGGGCGCGCGCGCGCGCGCGCGAGGCTTCACCCACAGCGCTGCGCAGCTACCTGCCCAA 440
|||||
Db 1521 CACGGTCAACGACACTGCGGGGAGCGGGGCTGTTGGGGCTGCTGTTGGCGCGCTGGG 1580
|||||
QY 441 CACGGTCAACGACACTGCGGGGAGCGGGGCTGTTGGGGCTGCTGCTGCGCGCTGGG 500
|||||
Db 1581 CGACAGCTGTGTTTACCTGCTGGCAGCTGCGCGCTCTTTTGTGCTGCTGCTGCCAG 1640
|||||
QY 501 CGACAGCTGTGTTTACCTGCTGGCAGCTGCGCGCTCTTTTGTGCTGCTGCTGCCAG 560
|||||
Db 1641 CTGCGCTTACAGGTGTGCGGGCGCGCTGTACAGCTGCGCGCTGCCACTCAGGCGCG 1700
|||||
QY 561 CTGCGCTTACAGGTGTGCGGGCGCGCTGTACAGCTGCGCGCTGCCACTCAGGCGCG 620
|||||
Db 1701 GCCCGCGCACACGCTAGTGGACCCCGAAGCGCTCTGGGATGCGAACGGGCTTGAACCA 1760
|||||
QY 621 GCCCGCGCACACGCTAGTGGACCCCGAAGCGCTCTGGGATGCGAACGGGCTTGAACCA 680
|||||
Db 1761 TAGGCTAGGAGCGCGGGTTCCTTGGCGCTTGGCAGCCCCGGGTGGAGAGGCGCGG 1820
|||||
QY 681 TAGGCTAGGAGCGCGGGTTCCTTGGCGCTTGGCAGCCCCGGGTGGAGAGGCGCGG 740
|||||
Db 1821 GGGCAGTCCAGCGAGTCTGCCGTTGCCCAAGAGCGCGCGCTGCGCTGCCCTGA 1880
|||||
QY 741 GGGCAGTCCAGCGAGTCTGCCGTTGCCCAAGAGCGCGCGCTGCGCTGCCCTGA 800
|||||

| | | | |
|----|------|--|------|
| Db | 12 | GTCTGCTGGGCACGTGGGAAGCCCTTGCGCCCGCGCCACCCCGCGGATGCCGCGCTCC | 71 |
| Qy | 22 | GTCTGCTGCGCACGTGGGAAGCCCTTGCGCCCGCGCCACCCCGCGGATGCCGCGCTCC | 81 |
| Db | 72 | CGTTCGCGAGCCGTGCGCTCCCTGCTGCGCAGCCACTACCGGAGGTGCTCCGCTGGCC | 131 |
| Qy | 82 | CGTTCGCGAGCCGTGCGCTCCCTGCTGCGCAGCCACTACCGGAGGTGCTCCGCTGGCC | 141 |
| Db | 132 | ACGTTCTGTCGGCGCCTGGGGGCCACAGGCTGGCGGCTGGTGCAGCGGGGACCCGGCG | 191 |
| Qy | 142 | ACGTTCTGTCGGCGCCTGGGGGCCACAGGCTGGCGGCTGGTGCAGCGGGGACCCGGCG | 201 |
| Db | 192 | GCTTTCGGCGCGCTGTGTGCCAGTGCCTGTGTGTGCTGCGTCCCTTGGAGCAGCGCGGCC | 251 |
| Qy | 202 | GCTTTCGGCGCGCTGTGTGCCAGTGCCTGTGTGTGCTGCGTCCCTTGGAGCAGCGCGGCC | 261 |
| Db | 252 | CCGCGCGCCCTTCCTTCGCGCAGGTGCTTCCCTGAAGAGCTGGTGGCCCGAGTGGTG | 311 |
| Qy | 262 | CCGCGCGCCCTTCCTTCGCGCAGGTGCTTCCCTGAAGAGCTGGTGGCCCGAGTGGTG | 321 |
| Db | 312 | CAGAGGCTGTGCAGCGCGCGCGAAGACGTGCTGGGCTTCGGCTTCGCGCTGCTGCAC | 371 |
| Qy | 322 | CAGAGGCTGTGCAGCGCGCGCGAAGACGTGCTGGGCTTCGCGCTTCGCGCTGCTGCAC | 381 |
| Db | 372 | GGGGCCCGGGGGGGCCCCCGAGGCCCTTTCACCAACAGCGTACCTGCCCAAC | 431 |
| Qy | 382 | GGGGCCCGGGGGGGCCCCCGAGGCCCTTTCACCAACAGCGTACCTGCCCAAC | 441 |
| Db | 432 | ACGGTGACACGACACTGCGGGGGAGCGGGGCTGGGGGCTGCTTGCGCCGCTGGGC | 491 |
| Qy | 442 | ACGGTGACACGACACTGCGGGGGAGCGGGGCTGGGGGCTGCTTGCGCCGCTGGGC | 501 |
| Db | 492 | GAGCAGCTGCTGCTTACCTGCTGGCAGCTCGCGCTCTTTGTGCTGGTGGCTCCAGC | 551 |
| Qy | 502 | GAGCAGCTGCTGCTTACCTGCTGGCAGCTCGCGCTCTTTGTGCTGGTGGCTCCAGC | 561 |
| Db | 552 | TGGGCTACACAGTGTGCGGGCGCGCTGTACAGCTCGCGCTGCGCTACGAGCCCGG | 611 |
| Qy | 562 | TGGGCTACACAGTGTGCGGGCGCGCTGTACAGCTCGCGCTGCGCTACGAGCCCGG | 621 |
| Db | 612 | CCCCGCCACACGCTAGTGGACCCCGAGGCTCTGGGATGCGAAGCGGCTGGAACCAT | 671 |
| Qy | 622 | CCCCGCCACACGCTAGTGGACCCCGAGGCTCTGGGATGCGAAGCGGCTGGAACCAT | 681 |
| Db | 672 | AGGCTCAGGAGGCGGGGCTCCCTTGGGCTTCAGAGCCCGGGTGCAGAGCGCGGG | 731 |
| Qy | 682 | AGGCTCAGGAGGCGGGGCTCCCTTGGGCTTCAGAGCCCGGGTGCAGAGCGCGGG | 741 |
| Db | 732 | GGCAGTGCCAGCGAGTCTGCCGTGCCCCAAGAGGCCAGGCGTGGCGCTGCCCTTGAG | 791 |
| Qy | 742 | GGCAGTGCCAGCGAGTCTGCCGTGCCCCAAGAGGCCAGGCGTGGCGCTGCCCTTGAG | 801 |
| Db | 792 | CCGAGCGGAGCGCCGTTGGGAGGGGTCTTGGGGCCACCCGGGACAGACGCTGGACCG | 851 |
| Qy | 802 | CCGAGCGGAGCGCCGTTGGGAGGGGTCTTGGGGCCACCCGGGACAGACGCTGGACCG | 861 |
| Db | 852 | AGTGACCGTGTTCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 911 |
| Qy | 862 | AGTGACCGTGTTCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 921 |
| Db | 912 | GAGGCTGCGCTCTCTGGCACGCGCCACTCCCAACCATCCGTGGGCGCCAGCACGCG | 971 |
| Qy | 922 | GAGGCTGCGCTCTCTGGCACGCGCCACTCCCAACCATCCGTGGGCGCCAGCACGCG | 981 |
| Db | 972 | GGCCCCCATCCACATCGGGGCGACACGCTCCCTGGGACAGCCTTGTCCCCCGGTGAC | 1031 |
| Qy | 982 | GGCCCCCATCCACATCGGGGCGACACGCTCCCTGGGACAGCCTTGTCCCCCGGTGAC | 1041 |
| Db | 1032 | GCGAGACCAAGCACTCTCTACTCTCTCAGCGGACAGGAGCAGCTCGCGCCCTCTTC | 1091 |
| Qy | 1042 | GCGAGACCAAGCACTCTCTACTCTCTCAGCGGACAGGAGCAGCTCGCGCCCTCTTC | 1101 |

| | | | |
|------------|--|---|------|
| Db | 1092 | CTACTCAGCTCTCTGAGGCCAGCTGACTGCGCTCGAGAGCTCGTGAGACCATCTTT | 1151 |
| Qy | 1102 | CTACTCAGCTCTCTGAGGCCAGCTGACTGCGCTCGAGAGCTCGTGAGACCATCTTT | 1161 |
| Db | 1152 | CTGGTTCCAGGCCCTTGGATGCCAGGACTCCCCGCAAGTTGCCCGCTCGCCACGGC | 1211 |
| Qy | 1162 | CTGGTTCCAGGCCCTTGGATGCCAGGACTCCCCGCAAGTTGCCCGCTCGCCACGGC | 1221 |
| Db | 1212 | TACTGGCAATTCGGGCCCTGTTTCTGAGAGCTGCTTGGGAACACAGCGCAGTGCCTTAC | 1271 |
| Qy | 1222 | TACTGGCAATTCGGGCCCTGTTTCTGAGAGCTGCTTGGGAACACAGCGCAGTGCCTTAC | 1281 |
| Db | 1272 | GGGTGCTCTCAAGACGACTGCCCGCTGCGAGTGGCGTCACCCCAGCAGCGGTGTC | 1331 |
| Qy | 1282 | GGGTGCTCTCAAGACGACTGCCCGCTGCGAGTGGCGTCACCCCAGCAGCGGTGTC | 1341 |
| Db | 1332 | TGTCCCGGAGAGGCCAGGCTCTGTGGCGGCCCCGAGGAGGAGACACAGACCCC | 1391 |
| Qy | 1342 | TGTCCCGGAGAGGCCAGGCTCTGTGGCGGCCCCGAGGAGGAGACACAGACCCC | 1401 |
| Db | 1392 | CGTCGCTGTGACGTGCTCCGCCAGCACAGCAGCCCTGGCAGGTGTACGGCTTCGTG | 1451 |
| Qy | 1402 | CGTCGCTGTGACGTGCTCCGCCAGCACAGCAGCCCTGGCAGGTGTACGGCTTCGTG | 1461 |
| Db | 1452 | CGGCCCTGCTCGCGCGGTGTGTGCGGCCCGAGGCCCTCGGGGTCCAGCACAGAACGC | 1511 |
| Qy | 1462 | CGGCCCTGCTCGCGCGGTGTGTGCGGCCCGAGGCCCTCGGGGTCCAGCACAGAACGC | 1521 |
| Db | 1512 | CGCTTCTCAGGACACCAAGAGTTTCATCTCCCTGGGAGAGCATGCCAAGCTCTCGCTG | 1571 |
| Qy | 1522 | CGCTTCTCAGGACACCAAGAGTTTCATCTCCCTGGGAGAGCATGCCAAGCTCTCGCTG | 1581 |
| Db | 1572 | CAGGAGCTGACGTGGAAGATGAGCGTGTGGGGCTGCGTGTGGCTGCGCAGGAGCCAGGG | 1631 |
| Qy | 1582 | CAGGAGCTGACGTGGAAGATGAGCGTGTGGGGCTGCGTGTGGCTGCGCAGGAGCCAGGG | 1641 |
| Db | 1632 | GTTGGCTGTGTTCCGGCCGCGAGCACCCTCTGCGTGGAGAGATCTGGCCAAAGTTCTTG | 1691 |
| Qy | 1642 | GTTGGCTGTGTTCCGGCCGCGAGCACCCTCTGCGTGGAGAGATCTGGCCAAAGTTCTTG | 1701 |
| Db | 1692 | CACGTGCTGATGATGTGTACGTCTGAGCTGCTCAGGTCTTTCTTTTATGTCAAGG | 1751 |
| Qy | 1702 | CACGTGCTGATGATGTGTACGTCTGAGCTGCTCAGGTCTTTCTTTTATGTCAAGG | 1761 |
| Db | 1752 | ACCACGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAA | 1811 |
| Qy | 1762 | ACCACGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAA | 1821 |
| Db | 1812 | AGCATTGGAATCAGACAGCACTTGAAGAGGTGTCAGCTGCGGAGCTGTCGGAAGCAGAG | 1871 |
| Qy | 1822 | AGCATTGGAATCAGACAGCACTTGAAGAGGTGTCAGCTGCGGAGCTGTCGGAAGCAGAG | 1881 |
| Db | 1872 | GTCAGGCAGCATCGGAGCCAGGCCCGCCCTGCTGACG | 1910 |
| Qy | 1882 | GTCAGGCAGCATCGGAGCCAGGCCCGCCCTGCTGACG | 1920 |
| RESULT | 4 | | |
| LOCUS | AF051911 | 3426 bp | mRNA |
| DEFINITION | Mus musculus telomerase reverse transcriptase mRNA, complete cds | | |
| ACCESSION | AF051911 | | |
| NID | g3005591 | | |
| KEYWORDS | house mouse. | | |
| SOURCE | Mus musculus | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| REFERENCE | 1 (bases 1 to 3426) | | |
| AUTHORS | Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R. | | |
| TITLE | Expression of mouse telomerase reverse transcriptase during development, differentiation, and proliferation | | |
| JOURNAL | Oncogene (1998) In press | | |
| REFERENCE | 2 (bases 1 to 3426) | | |

| | | | | | |
|------------|---|--|------|-----|-------------|
| RESULT | 4 | | | | |
| LOCUS | AF051911 | 3426 bp | mRNA | ROD | 02-APR-1998 |
| DEFINITION | Mus musculus telomerase reverse transcriptase mRNA, complete cds. | | | | |
| ACCESSION | AF051911 | | | | |
| NID | 93005591 | | | | |
| KEYWORDS | | house mouse. | | | |
| SOURCE | | Mus musculus | | | |
| ORGANISM | | Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| REFERENCE | 1 | (bases 1 to 3426) | | | |
| AUTHORS | Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A. | | | | |
| TITLE | Expression of mouse telomerase reverse transcriptase during development, differentiation, and proliferation | | | | |
| JOURNAL | Oncogene (1998) | In press | | | |
| REFERENCE | 2 | (bases 1 to 3426) | | | |

AUTHORS Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and Depinho, R.A.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-1998) Microbiology and Immunology, Albert
Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY
10461, USA

FEATURES

source

1. 3426

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="13"

/map="MGD offset position 40; 40 cM from centromere"

30. 3398

/note="mTERT"

/codon_start=1

/product="telomerase reverse transcriptase"

/db_xref="pid:g300592"

/translation="MTRAPRCPAVRSLRSRYREVWPLATFVRRLGPEGRRLLVQPGDP

KIYRLVAOCLVCHMWSQPPADLSHOVSSSLKELVARVVRICERNERNVLAFGPE

LLNEARGPPNPAFTSSVRSYLPNTVETLRYSGAWMLLSRVGDDLVIYLLAHCLALYL

LVPSCAYQCGSLYQICATDIIWPSVASIRTPRPVGRNFTNRLQIQLSSROE

APKPLALPRTGTRKRLSLTSVPSAKARCIYPRVEEGPHRQVLPFTPSKSWPSP

ARSPEVTAEDLSKSKVSDLSGVSCKKPSSTLSLSPRONAFQRLPFIETRW

FLYRGDQERLNSFLSNLPNLTGARLVEIFLGSRPTSGPLCRHLRSLRRYW

QMRPLFOCLLVNAECOVRLRLSHCRFTANQVTDALNTSPHMLDLRLHSSPWQ

VYGLFRALCKVVSASLWGRHNRERFKLKEISLGYKLSLQELMKMKVEDCH

WLRSPGKDRVPAEHLRLERILATFLWMDTVVQLRSFFIITESTQKNRLFY

RKSVMKLSIGVROHLERVRLELSOEVEVRHDDTWLAMPICRLRIPKPNLRPIV

NMSYGTALRGRKQAHFTQRLKTLFSLNLYERTKPHLMGSSVGLMNDIYVTRWA

FVLVRALDQTPRMVFKVADYTDAYDAIPOGLKVEVWAMIRHSESYTCROYAVRR

DSOGVHKSPFRVYTLSDLOPYMGQFLKHLQSDASALNSVVEIOSISNMSSESL

FDFLFLHRSVVKIDRCYTQCOGIPQGSLSLLSLCLSGEDMENKLFARVQDGLL

LRFVDDFLVYPLHDQAKFTLSLVHVGPEGMINKQIVNFPFPIGLGAAPIQ

LPACHLFPWCGLLDLTQLEVFCDYSGTAQTSIKTSITFOSFVAGKTRMKLLSVLR

LKHGFLDVLVNSLOTVGINIYKIFLLQAYRFHACVIQLPFDQRVKNLTFTGLIIS

SOASCRAILKVNKPTLKASGFPPEAAHLWCYQAFLLKLAHSVIYKCLLGLPLRT

AQKLKRLPEATVILKAAADPALSTDQFILD"

BASE COUNT 746 a 979 c 907 g 794 t

ORIGIN

Query Match 23.6%; Score 453; DB 28; Length 3426;

Best Local Similarity 64.3%; Pred. No. 0.00e+00;

Matches 1216; Conservative 0; Mismatches 592; Indels 84; Gaps 3;

Db 14 CCGCGCTTGACACAAATGACCCGCGCTCTGCTGGCCCGCGGTGCGCTCTCTCGCG 73

Qy 51 CCGCGCAACCCCGCGATGCGCGCGCTCCCGCTCGCGAGCGCTCCCTCGCTGCG 110

Db 74 CAGCGATACCGGAGGTGTGCGCGTGGCAACCTTGTGGCGCGCTGGGCGCGAGGG 133

Qy 111 CAGCCACTACCGAGGTGTGCGCGTGGCAACCTTGTGGCGCGCTGGGCGCGAGGG 170

Db 134 CAGCGCGCTTGCAACCCCGGAGCCGGAAGATCTACCGCACTTTGGTTGCCCAATGCGT 193

Qy 171 CTGCGCGCTGGTGACCGCGGAGCCCGCGCGCTTTCCCGCGCTGGTGGCCAGTGCCT 230

Db 194 AGTGTGATGCTAGTGGGCTACAGCTCCACCTCGCGACCTTCTCCACCAAGGTGTC 253

Qy 231 GGTGTGCGTGGCTGGGAGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 290

Db 254 ATCCCTGAAAGAGTGTGGCGAGGTTGTGACAGAGCTCTCGAGCGCAACGAGAGAA 313

Qy 291 CTGCGCTGAAGAGCTGTGGCGCGAGTGTGCGAGAGCTGTGCGAGCGCGCGGAGAA 350

Db 314 CCGTGTGCTTTTGGCTTTGAGCTGCTTAACAGAGCCGAGGCGGCGCTCCCATGGCCTT 373

Qy 351 CCGTGTGCGCTTCCGCTTCCGCTGCTGGACGGGCGCGCGGCGCGCGCGCGCGCTT 410

Db 374 CACTAGTACCGTGTAGTACTTCCCAACACTGCTTATAGACCCCTGCGTGTGAGTGG 433

Qy 411 CACACACCGTGTGCGAGCTACTTCCCAACACTGCTTCCCAACACTGCGGCGGAGGG 470

Db 434 TGCATGGATGCTACTTGTGAGCGGAGTGGCGGACGACCTGCTGCTACCTGCTGSCACA 493

Qy 471 GGCCTGGGGGCTGCTGCTGCGCGCGCTGGCGGACAGCTGCTGCTTACCTGCTGGCAG 530

Db 494 CTGTGCTCTTTATCTTCTGTGTCGCCCGCAGCTGTGCTTACAGCTGTGTGGTCTCCCT 553

Qy 531 CTGCGCGCTCTTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590

Db 554 GTACCAATTTGTGCCACCGGATATCTGCGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCT 613

Qy 591 GTACCAAGCTGGCGCTGCCACTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 650

Db 614 CCGACCGCTGGCGAGGAATTTCACTAACCCTTAGCTTCTTACCAACAGATCAAGACGAGT 673

Qy 651 CGCTCTGGGATCG-----AACGGCGCTGGAACCATAG 683

Db 674 TCGCCAGGAACGACCGAAACCCCTGGCTTGGCCATCTCAGGTTACAAAGAGGATCTGAG 733

Qy 684 CGTCAGGAGCGCGGGTCCCTTGGCTTGGCCAGCGCGCGCGCGCGCGCGCGCGCGGG 743

Db 734 TCTCAGCAGTACAAGTGTGCTTCACTAAGAGCGCCAGATGCTATCTGCTCCCGAGAT 793

Qy 744 CAGTGCACCGGAAGTCTGCGCTTGGCCAAAGAGCGCGCGCGCGCGCGCGCGCGCG 803

Db 794 GGAGGAGGAGCCACACAGGAGGTGCTACCAACCCCTCAGGCAAAATCATGGGTGCCAAG 853

Qy 804 GGAGCGAGCGCGCTTGGCGAGGGTCTTGGCGCGCGCGCGCGCGCGCGCGCGCGAG 863

Db 854 TCCTGCTCGGTCCCGAGGTGCTTACTGCGAGAGAAAGATTTGCTTCTTAAGGAAAGGT 913

Qy 864 TGACCGGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 923

Db 914 GTCTGACCTGAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961

Qy 924 GGGTGCCTCTCTGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 983

Db 962 CTCCACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1021

Qy 984 CCCCCCTCCACATCG 1043

Db 1022 GACGAGCATTTCTTACTTCCAGGGAGATGGCCAAAGAGCGTCTAAACCCCTCATTCCT 1081

Qy 1044 CGAGACCAAGCACCTTCTCTTACTTCTCAGCGGACAAAGAGCGAGCTGCGCGCGCT 1103

Db 1082 ACTCAGCAACCTCAGCGCTTACTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1141

Qy 1104 ACTCAGCTCTCTGAGCG 1163

Db 1142 GGGCTCAAGGCTAGGACATCAGGACCACTCTGCGAGGACACCGCTCTATCGCGTCGATA 1201

Qy 1164 GGGTCCAGGCGCTGATGCCAGGACTTCCCGCAGGTTGCGCGCGCGCGCGCGCGCG 1223

Db 1202 CTGCGAGATGCGCGCGCTTGTTCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1261

Qy 1224 CTGCGAAATGCGCGCGCTGTTCTTCTGAGCTGCTTGGGAAACCGCGAGTGCCTTAC 1283

Db 1262 CAGACTCTCAGTTCATTTGAGGTTTGGAAACAGCAACCAACAGGTTGACAGATGCTT 1321

Qy 1284 GGTGCTCTCTAAGACGACCTGCGCGCTGCGAGTCTGCGCTGCGCTGCGCTGCGCTG 1343

Db 1322 -----GAACACCG 1336

Qy 1344 TGCCCGGAGGAGCG 1403

Db 1337 GCACCTCATGATTTGCTTCCCGCTGCGAGCTGCTGCGAGTATATGTTTCTTCTG 1396

Qy 1404 TCGCTGTGTCAGCTGCTCGCGACGACAGCGCGCTGCGAGTGTACGGCTTCTGCTG 1463

Db 1397 GGCCTGCTCTGCAAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1456

Qy 1464 GGCCTGCTGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1523

Db 1457 CTTCTTTAAGAACTTAAGAGTTTCATCTGTTGGGAAATACGCGCAAGCTTATCCTGCA 1516

Qy 1524 CTTCTCAGGAGACCAAGAAATTCATCTCTCTGCGGAGGAGTCCAGGCTCTCGCTGCA 1583

Db 949 TCACACCCCGCCAAATGCTTTTTCAGCTCAGCGCATTTATTGACACAGACATTTTCCTT 1008
QY 1000 CGCGCCACACGCTCCCTGGGACACGCTTGTCCCGGTGTAGTCGCGGACCAAGCACTTC 1059
Db 1009 TACTCCAGGGAGATGGCCCAAGAGCTCTAAACCCCTCATTTCTTACTACTCAGCAACCTCCAG 1068
QY 1060 CTCTACTCTCAGGCGACAAGAGAGAGCTGCGGCCCTCTTCTTACTCAGCTCTCTGAGG 1119
Db 1069 CTTAACTTCTGCTGGGCGCCAGGAGACTGCTGGAGATCATTTTCTTGGGCTCAAGGCTTAGG 1128
QY 1120 CCCAGCTGACTGGGCTCGGAGGCTCGTGGAGACATCTTCTTGGTTCAGAGCCCTGG 1179
Db 1129 ACATCAGGACCACTCTGCAGGACACACCTCTATGCGTGCATATGCGAGATGCGGCC 1188
QY 1180 ATGCGAGGACTCCCGCAGGTGGCCCGCTGCCCGCTGCTGCAATGCGGCC 1239
Db 1189 CTTTCCAACTGCTGCTGAACCATGAGAGTGCAGATGCGCATATGTCAGACTCTCAGGTCA 1248
QY 1240 CTTTCTTGAGCTGCTTGGGAACACCGCGAGTGCCTTACGGGGTCTCTCAAGAG 1299
Db 1249 CATTCAGGTTTTCGAACCAACCAACAGGTGACAGATGCTT----- 1292
QY 1300 CACTGCCCTCGAGCTCGCTCACCCAGCAGCGGTGTCTGTGCCCGGAGAGGCC 1359
Db 1293 -----GAACACAGCCCAACCGCACTCATGATTG 1323
QY 1360 CAGGCTCTGTGGCGCCCGCCGAGGAGGAGACACAGACCCCGCTGCTGCAAGT 1419
Db 1324 CTCGCGCTCAGCAGAGTCCCTGGGAGTATATGTTTCTTGGGCTGCTCTGCAAG 1383
QY 1420 CTCGCGCAGCAGACGCCCCCTGGCAGGTGTACGGCTTCTGCGGGCTGCTGCGCGG 1479
Db 1384 GTGGTCTCTAGTCTTGGGTTACAGGCACATAGCGCGCTTCTTTAAGAACTTA 1443
QY 1480 GTGGTCCCGCCAGCTCTGGGCTCCAGGCACACAGACGCGCTTCTCAGGACACC 1539
Db 1444 AAGAAGTTATCTGTTGGGAAATACGGCAAGCTATCACTGCAGAACTGATGTGAAG 1503
QY 1540 AAGAAGTTATCTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGAGTGAAG 1599
Db 1504 ATGAAGTAGAGATTGCCATGCTGCTGCGCAGCAGCGCGGGAAGACCGTCTCCCGCT 1563
QY 1600 ATGAGGCTCGGGAGTGGCTTGGCTGCGCAGGAGCCAGGGGTGGCTGTCTCCGCC 1659
Db 1564 GCAGACACCGCTCTGAGGAGGAGTCTGCTGCTAGCTTCTTCTTGGCTGATGGACACA 1623
QY 1660 GCAGACACCGCTCTGCTGAGGAGATCTGCGCAAGTCTCTGCACTGCTGATGATG 1719
Db 1624 TAGTGGTACAGCTGTTAGGTCTATTTTATCATCAGACAGAGACATTCAGAGAAGAC 1683
QY 1720 TAGTCTGCTGAGCTGCTCAGGTCTTTCTTTTATGTACGAGACACCGCTTCAAAGAAC 1779
Db 1684 AGGCTCTTCTTCTACGTAAGAGTGTGTGAGCAAGCTCAGAGCATTTGAGTCAAGGAA 1743
QY 1780 AGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTCGAAGCATTTGAATCAGACAG 1839
Db 1744 CACCTTGAGAGGCTGCGCTACGGAGTGTGTCAAGAGGAGGTGAGGATCAGCAGGAC 1803
QY 1840 CACTTGAAGAGGTCAGCTGCGGGAGCTGTGCGAAGCAGAGGTGAGCAGCATCGGNA 1899
Db 1804 ACTGCTAGCATGC 1819
QY 1900 GCCAGCGCGCCCTGC 1915

RESULT 6
LOCUS AF029235 201 bp mRNA ROD 12-NOV-1997
DEFINITION Mus musculus telomerase catalytic subunit mRNA, partial cds.
ACCESSION AF029235
NID 92605902
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus

Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 201)
AUTHORS Drissi,R. and Cleveland,J.L.
TITLE Partial sequence of Mus musculus telomerase catalytic subunit homolog
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 201)
AUTHORS Drissi,R. and Cleveland,J.L.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1997) Biochemistry, St Jude Children's Research Hospital, 322 North Lauderdale, Memphis, TN 38105, USA
FEATURES
source 1..201
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="32D.3"
/codon_start=1
/product="telomerase catalytic subunit"
/db_xref="PID:g2605903"
/translation="FFVYTESTFKNRLFFYRKSVMKLSQIGVHQHLRVRLRELQ
EEVHHQDTWLAMPICELRIPK"
BASE COUNT 51 a 54 c 55 g 41 t
ORIGIN
Query Match 5.7%; Score 110; DB 28; Length 201;
Best Local Similarity 82.0%; Pred. No. 3,62e-55;
Matches 141; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
Db 1 TTTCTTTATGTCAGGAGAGCACAATCCAGAAACAGGCTTCTTCTTACCGTAAGAGT 60
QY 1744 TTTCTTTATGTCAGGAGAGCACCAGTTCCTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGT 1803
Db 61 GTGTGGACCAAGCTGTCAGAGCATTTGGAGTCAGGCAACACCTTGAGAGAGTCGGCTACGG 120
QY 1804 GTCTGGAGCAAGTTGCAAGCATTTGGAATCAGACAGCACTTTGAAGAGGTCGACGTCGG 1863
Db 121 GAGCTGTCAAGAGGAGGTGAGCATCAGCAGCAGCAGCAGTGGCTAGCCATGC 172
QY 1864 GAGCTGTGAGGAGGAGGTGAGCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1915
RESULT 7
LOCUS I66494 7218 bp DNA PAT 23-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION I66494
NID 92724471
KEYWORDS .
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7218)
AUTHORS Dornier,F., Scheifflinger,F. and Falkner,F.Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
source 1..7218
/organism="unknown"
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN
Query Match 3.5%; Score 67; DB 21; Length 7218;
Best Local Similarity 2.4%; Pred. No. 6,05e-25;
Matches 9; Conservative 211; Mismatches 153; Indels 0; Gaps 0;
Db 1055 GGAGCTGCGATYYY 1114
QY 921 GGAGGTCGCTCTCTGCGACGCGCACTCCACCCATCCGTCGCGCCGACGACGCG 980
Db 1115 YY 1174
QY 981 GGGCCCCCATCCACATCGCGGCCACCACTGCTCCCTGGGACACGCTTGTCCCGGTGTA 1040


```

Db 1175 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1234
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1041 CGCCGAGACCAAGCATTCTCTACTCTCAGCGACAAGGAGCAGCTCGCGCCCTCCTT 1100
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1235 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1294
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1101 CCTACTAGCTCTCTGAGCCGACCTGACTGGCGCTCGGAGGCTCGGGAGACCATCTT 1160
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1295 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1354
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1161 TCTGGTTCAGGCCCTGGATGCCAGGACATCCCGCAGGTTGCCCGCTGCCCGCAGCG 1220
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1355 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1414
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1221 CTACTGGCAATCGGCCCTCTTCTGAGCTGCTTGGGAACCAACGCGAGTGCCTCA 1280
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1415 YYYYYYYYYYYY 1427
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1281 CGGGTGCTCTC 1293
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8 I66494 7218 bp DNA PAT 23-DEC-1997
LOCUS Sequence 14 from patent US 5670367.
DEFINITION I66494
ACCESSION 92724471
NID
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Location/Qualifiers
source
1..7218
/organism="unknown"
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN

Query Match 2.3%; Score 44; DB 21; Length 7218;
Best Local Similarity 0.8%; Pred. No. 4.31e-10;
Matches 3; Conservative 196; Mismatches 155; Indels 0; Gaps 0;

Db 1062 GCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1121
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 1910 GCGGCCCTGGCTCCCGATGCTGCTGACCTGCTTCCGACAGTCCGCGAGCTGCACC 1851
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1122 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1181
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 1850 CTCTCAAGTGTGCTGATTCATGCTTGTCAACTGTCTCCAGACACTCTCTCGGTAG 1791
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1182 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1241
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 1790 AAAAGAGCCCTGCTTTTGAACGTGCTCCGTCGACATAAAGAAAGACCTGAGCAGC 1731
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1242 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1301
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 1730 TCGAGCAGCTACACTATCAGCAGCTGAGGACTTGGCCAGGATCTCTCAGCAGCA 1671
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1302 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1361
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 1670 CGGTGCTCGCGCGGAACAGCAGCAACCCCTGGCTCTCGCAGCAGCAAGCGCAGTCC 1611
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1362 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1415
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 1610 CGCAGGCTCATCTCCAGCTCAGCTCCTCGCAGCAGAGCTTGGCATGCTTCCCC 1557
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9 AC005369 74371 bp DNA PRI 01-AUG-1998
LOCUS Homo sapiens chromosome 5, BAC clone 119j3 (LBNL H175), complete
DEFINITION

```

```

sequence.
AC005369
NID 93367505
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 74371)
Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
Sequencing of human chromosome 5
Unpublished
2 (bases 1 to 74371)
Ricke,D.O.
Large Scale Sequence Analysis and Annotation with the Sequence
Comparison Analysis (SCAN) System
Unpublished
3 (bases 1 to 74371)
Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
Direct Submission
Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
Sequence submitted by:
DOE Joint Genome Institute.
FEATURES
Location/Qualifiers
source
1..74371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="5q"
/clone="119j3"
/chromosome="5"
/note="LBNL H175"
893..1030
/rpt_family="Alu"
2295..2438
/rpt_family="Alu"
2818..2859
/note="(G1)21"
/rpt_type=tandem
join(3246..3410,3721..3828)
/standard_name="RLF"
/note="65% & 69% protein identity GenPept:U22377"
3431..3724
/rpt_family="Alu"
3707..3728
/note="(A)22"
/rpt_type=tandem
/rpt_unit=A
4366..4661
/rpt_family="Alu"
5327..5602
/rpt_family="Alu"
6586..6956
/rpt_family="L1"
6647..6684
/note="(CA)19"
/rpt_type=tandem
/rpt_unit=CA
7113..7373
/rpt_family="Alu"
complement(7830..8185)
/standard_name="possible repeat"
8258..8503
/rpt_family="Alu"
9070..9387
/rpt_family="Alu"
complement(9740..9845)

```



```
KEYWORDS      Unknown.
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 215)
AUTHORS        Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
TITLE          Plant inhibitors of fungal polygalacturonases and their use to
JOURNAL        control fungal disease
FEATURES       Patent: US 5569830-A 5 29-OCT-1996;
               Location/Qualifiers
               1..215
               /organism="unknown"
BASE COUNT    15 a      8 c      25 g      26 t      141 others
ORIGIN

Query Match      1.8%; Score 34; DB 21; Length 215;
Best Local Similarity 12.9%; Pred. No. 2,93e-04;
Matches 17; Conservative 56; Mismatches 59; Indels 0; Gaps 0;

Db 83 HTNVSGADSKVTDYSNAGTSSSSNGGTDGNNRSGADSYGSSKTAITSNRRTGKTANNAVD 142
Qy 130 CTGGCGCTGGCCACGTCCTGTCGGCGCTGCGGCCGCCAGGGCTGCGGCTGGTCGAGCGC 189
Db 143 SRNMGDASVGSNDKNTKHKAKNSADGKVGSKNNGDRNNRYGTGKSNVSNCCGGGNKRDVS 202
Qy 190 GGGGACCGGGCGGCTTCGCGCGCTGTGGTGGCCCGAGTGCCTGTGTGCTGCCCTGGGAC 249
Db 203 SYANNKCGSSC 214
Qy 250 GCACGGCGGCC 261

RESULT 11
LOCUS      128278      215 bp      DNA      PAT      30-OCT-1996
DEFINITION Sequence 5 from patent US 5569830.
ACCESSION 128278
NID        91819054
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 215)
AUTHORS    Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
TITLE      Plant inhibitors of fungal polygalacturonases and their use to
JOURNAL    control fungal disease
FEATURES   Patent: US 5569830-A 5 29-OCT-1996;
           Location/Qualifiers
           1..215
           /organism="unknown"
BASE COUNT 15 a      8 c      25 g      26 t      141 others
ORIGIN

Query Match      1.8%; Score 35; DB 21; Length 215;
Best Local Similarity 13.2%; Pred. No. 8,13e-05;
Matches 26; Conservative 76; Mismatches 95; Indels 0; Gaps 0;

Db 6 MSSSVVSRATSCNDKAKKDGNTTSWTDCCNRTWGYCDTDTTYRVNDSGHNKYSSAN 65
Cp 1615 AGTCCCGCAGCGTCATCTCCACGTCAGCTCCTGCAGCGAGAGCTTGGCATCTCCCCA 1556
Db 66 YNYGNNYGAAKTHYITHTNVSGADSKVTDYSNAGTSSSSNGGTDGNNRSGADSYGSSKT 125
Cp 1555 GGGAGATGAACCTCTTGTGTTCTCTGAGGAAGCGGCTGCTGTGCTGGAGCCCCAGA 1496
Db 126 AMTNRRTGKTANNADVSRNMGDASVGSNDKNTKHKAKNSADGKVGSKNNGDRNNRYGTGT 185
Cp 1495 GGCCTGGGGGACACCGCGCGGCGAGCGGCCGCCAGGACCGGTACACCTGCCAGGGGC 1436
Db 186 KSNVSNCCGGGNKRDVS 202
Cp 1435 TGCTGTGCTGGCGGAGC 1419
```

```
RESULT 12
LOCUS      AC005369      74371 bp      DNA      PRI      01-AUG-1998
DEFINITION Homo sapiens chromosome 5, BAC clone 119j3 (LBNL H175), complete
sequence.
ACCESSION  AC005369
NID         G3367505
KEYWORDS    HTG.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 74371)
            Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
            Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M.,
            Rojeski,H., Subramanian,S. and Martin,C.H.
            Sequencing of human chromosome 5
            Unpublished
            2 (bases 1 to 74371)
            Ricke,D.O.
            Large Scale Sequence Analysis and Annotation with the Sequence
            Comparison Analysis (SCAN) System
            Unpublished
            3 (bases 1 to 74371)
            Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
            Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M.,
            Rojeski,H., Subramanian,S. and Martin,C.H.
            Direct Submission
            Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome
            Institute, Lawrence Berkeley National Laboratory, MS 74-157,
            Berkeley, CA 94720, U.S.A.
            Sequence submitted by:
            DOE Joint Genome Institute.
            Location/Qualifiers
            1..74371
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /map="5q"
            /clone="119j3"
            /chromosome="5"
            /note="LBNL H175"
            893..1030
            /rpt_family="Alu"
            2295..2438
            /rpt_family="Alu"
            2818..2859
            /note="(GT)21"
            /rpt_type=tandem
            /rpt_unit=GT
            join(3246..3410,3721..3828)
            /standard_name="RLF"
            /note="65% & 69% protein identity GenPept:U22377"
            3431..3724
            /rpt_family="Alu"
            3707..3728
            /note="(A)22"
            /rpt_type=tandem
            /rpt_unit=A
            4366..4661
            /rpt_family="Alu"
            5327..5602
            /rpt_family="Alu"
            6586..6956
            /rpt_family="L1"
            6647..6684
            /note="(CA)19"
            /rpt_type=tandem
            /rpt_unit=CA
            7113..7373
            /rpt_family="Alu"
            complement(7830..8185)
            /standard_name="possible repeat"
            8258..8503
            /rpt_family="Alu"
            repeat_region
```


[illegible]

Search completed: Thu Dec 24 13:04:28 1998
Job time : 12050 secs.

THIS PAGE BLANK (USPTO)

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 13:04:47 1998; MasPar time 5213.74 Seconds
Tabular output not generated. 659.468 Million cell updates/sec

Title: >US-08-951-733-13

Description: (1-1920) from US08951733.seq (1 of 2)

Perfect Score: 1920

N.A. Sequence: 1 CAGCGCTCCGGGAGCGCTG.....CCAGCGCCGCCCTGCTGACG 1920
Comp: GTGCGCAGCGCCGTCGGAC.....GGTCCGGCGGCGGACGACTGC

Scoring table: TABLE default

Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs; 895388244 bases x.2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: emb1-est55

Database: 1:em_est1 2:em_gss1 3:em_gss2 4:em_gss3

5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est20 17:gb_est21
18:gb_est22 19:gb_est23 20:gb_est24 21:gb_est25 22:gb_est26
23:gb_est27 24:gb_est28 25:gb_est29 26:gb_gss1 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 11.357; Variance 2.736; scale 4.151

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|---------------------------------|-----------|
| 1 | 227 | 11.8 | 389 | 8 | AA281296 zt08g02.r1 NCI_CGAP_GC | 0.00e+00 |
| 2 | 58 | 3.0 | 252 | 12 | AA754459 97SN1787 Rice Immature | 1.27e-46 |
| 3 | 54 | 2.8 | 252 | 12 | AA754459 97SN1787 Rice Immature | 7.79e-41 |
| 4 | 49 | 2.6 | 247 | 12 | AA754458 97SN1784 Rice Immature | 9.33e-34 |
| 5 | 44 | 2.3 | 247 | 12 | AA754458 97SN1784 Rice Immature | 6.93e-27 |
| 6 | 31 | 1.6 | 660 | 11 | AF034177 Homo sapiens ntcon6 co | 2.10e-10 |
| 7 | 28 | 1.5 | 1287 | 12 | AF038250 Homo sapiens clone ntc | 5.22e-07 |
| 8 | 28 | 1.5 | 1287 | 12 | AF038250 Homo sapiens clone ntc | 5.22e-07 |
| 9 | 27 | 1.4 | 317 | 27 | A0008007 CIT-HSP-2288C3.TF CIT- | 6.33e-06 |
| 10 | 26 | 1.4 | 400 | 16 | H44707 YP24c08.r1 Homo sapien | 7.23e-05 |
| 11 | 27 | 1.4 | 660 | 11 | AF034177 Homo sapiens ntcon6 co | 6.33e-06 |
| 12 | 26 | 1.4 | 2275 | 11 | AF034173 Homo sapiens ntcon2 co | 7.23e-05 |

| C | 13 | 25 | 143 | 18 | AI171394 | EST217350 Normalized r | 7.73e-04 |
|---|----|----|------|----|----------|-------------------------|----------|
| | 14 | 25 | 181 | 16 | AA386387 | EST81369 Prostate glan. | 7.73e-04 |
| | 15 | 25 | 187 | 13 | AA855630 | vw8h11.r1 Stratagene | 7.73e-04 |
| | 16 | 25 | 213 | 16 | R72797 | Y109c09.r1 Homo sapien | 7.73e-04 |
| | 17 | 25 | 242 | 15 | AI012107 | EST206558 Normalized r | 7.73e-04 |
| | 18 | 25 | 278 | 13 | AA882818 | TENS0393 T. cruzi epim | 7.73e-04 |
| | 19 | 25 | 302 | 15 | AI010844 | EST205295 Normalized r | 7.73e-04 |
| | 20 | 25 | 318 | 16 | H50134 | yo27a07.r1 Homo sapien | 7.73e-04 |
| | 21 | 25 | 336 | 18 | AI172396 | EST218402 Normalized r | 7.73e-04 |
| | 22 | 25 | 359 | 24 | AA303595 | EST16305 Aorta endothe | 7.73e-04 |
| | 23 | 25 | 382 | 5 | R54656 | YJ74d04.r1 Homo sapien | 7.73e-04 |
| | 24 | 25 | 420 | 17 | AI071447 | UI-R-C1-ku-c-05-0-UI.s | 7.73e-04 |
| | 25 | 25 | 438 | 14 | AA924206 | UI-R-11-dx-a-11-0-UI.s | 7.73e-04 |
| | 26 | 25 | 478 | 14 | AA923372 | UI-R-11-ee-d-07-0-UI.s | 7.73e-04 |
| | 27 | 25 | 487 | 16 | R86860 | ym86a03.r1 Homo sapien | 7.73e-04 |
| | 28 | 25 | 504 | 15 | AI010120 | EST204571 Normalized r | 7.73e-04 |
| | 29 | 25 | 511 | 15 | AA997735 | UI-R-C2-hn-b-08-0-UI.s | 7.73e-04 |
| | 30 | 25 | 511 | 17 | AI072026 | UI-R-C2-nd-e-12-0-UI.s | 7.73e-04 |
| | 31 | 25 | 540 | 14 | AA901048 | UI-R-11-dv-e-09-0-UI.s | 7.73e-04 |
| | 32 | 25 | 590 | 10 | AA622227 | no3c05.sl NCI_CGAP.Pr | 7.73e-04 |
| | 33 | 25 | 677 | 18 | AI105428 | EST214717 Normalized r | 7.73e-04 |
| | 34 | 25 | 705 | 14 | AA941159 | LD25143.5prime LD Dros | 7.73e-04 |
| | 35 | 24 | 275 | 14 | AA899633 | UI-R-20-dg-f-09-0-UI.s | 7.71e-03 |
| | 36 | 24 | 288 | 23 | AA077661 | 7B35F02 Chromosome 7 F | 7.71e-03 |
| | 37 | 24 | 394 | 16 | R92606 | yq07g02.sl Homo sapien | 7.71e-03 |
| | 38 | 24 | 398 | 5 | TA7934 | ybi8d04.sl Homo sapien | 7.71e-03 |
| | 39 | 24 | 426 | 9 | AA591041 | vl47b11.sl Stratagene | 7.71e-03 |
| | 40 | 24 | 456 | 6 | C19919 | Rice cDNA, partial seq | 7.71e-03 |
| | 41 | 24 | 496 | 23 | AA243831 | zr67b09.r1 Soares NHM | 7.71e-03 |
| | 42 | 24 | 496 | 10 | AA429394 | zw32h12.sl Soares ovar | 7.71e-03 |
| | 43 | 24 | 550 | 7 | AA495273 | fa02e07.r1 Zebrafish i | 7.71e-03 |
| | 44 | 24 | 607 | 18 | AI170655 | EST216590 Normalized r | 7.71e-03 |
| | 45 | 24 | 2275 | 11 | AF034173 | Homo sapiens ntcon2 co | 7.71e-03 |

ALIGNMENTS

| RESULT | 1 |
|------------|---|
| LOCUS | AA281296 389 bp mRNA EST 14-AUG-1997 |
| DEFINITION | zt08g02.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:712562 5', mRNA sequence. |
| ACCESSION | AA281296 |
| NID | g1924194 |
| KEYWORDS | EST. |
| SOURCE | human. |
| ORGANISM | Homo sapiens |
| REFERENCE | 1 (bases 1 to 389) |
| AUTHORS | NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. |
| TITLE | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index |
| JOURNAL | Unpublished (1997) |
| COMMENT | Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 2187 Std Error: 0.00 Seq primer: -28m13 rev2 ET from Amersham High quality sequence stop: 385. |

FEATURES

1. 389
Location/Qualifiers
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD+), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer

THIS PAGE BLANK (USPTO)

Best Local Similarity 20.9%; Pred. No. 3.86e-09;
Matches 31; Conservative 57; Mismatches 60; Indels 0; Gaps 0;

Db 45 bcyrraggnccggggyccgagcycaayycdchvgccgymrttthhyrmnbnyrd 104
QY 829 TCTGGGCCACCCGGGAGGACGGCTGTGGAGGCTGCTCTGGGACCGGCCAC 888
Db 105 ynrsdaaayccyrrsvkydcncynachdhdyvvybbbyvnhnnncncbnnhvhcnv 164
QY 889 CTTGGCAGACCCGCGAAGACCACTTGTGGAGGCTGCTCTGGGACCGGCCAC 948
Db 165 hbnhnwayrhrdardddvhcvcgcc 192
QY 949 TCCACCCATCCGTGGCGCGCAGCACC 976

RESULT 4

| ID | Q70466 standard; DNA; 114 BP. |
|----|---|
| ID | Q70466; |
| AC | Q70466; |
| DT | 05-APR-1995 (first entry) |
| DE | Generic DNA sequence to generate a random TSAR-9 peptide library. |
| KW | TSAR: totally synthetic affinity reagent; synthetic; binding domain; |
| KW | effector domain; concatenated heterofunctional protein; linker; |
| KW | direct; rapid; detection; screening; treatment; generic; ss. |
| OS | Synthetic. |
| FS | |
| FH | Key |
| FT | Location/Qualifiers |
| FT | misc_feature |
| FT | 55..60 |
| FT | /*tag= a |
| FT | /note= "this sequence represents 'Z'; Z can be a |
| FT | sequence of 6, 9 or 12 nucleotides (see |
| FT | comments)" |
| FT | |
| PN | WO9418318-A. |
| PD | 18-AUG-1994. |
| PF | 01-FEB-1994; U00977. |
| PF | 01-FEB-1994; US-013416. |
| PR | 30-DEC-1993; US-176500. |
| PR | 31-JAN-1994; US-189331. |
| PI | (UYNC-) UNIV NORTH CAROLINA. |
| PA | Fowlkes DM, Kay BK; |
| DR | WPI; 94-279739/34. |
| DR | P-PSDB; R65152. |
| PT | Identifying proteins or peptide(s) which bind a ligand - by |
| PT | screening a recombinant vector library expressing fusion proteins |
| PT | comprising a binding domain and an effector domain |
| PS | Disclosure; Page 35: 255pp; English. |
| CC | Q70466 is a generic DNA sequence used to generate random TSAR (Totally |
| CC | Synthetic Affinity Reagents) peptides. This generic formula can also be |
| CC | represented as follows: (XNNB)(10(TGC)2(NNB)42(NNB)8(TGC)(NNB) |
| CC | -9Y. X and Y are flanking restriction sites (X is not the same as Y) |
| CC | that are not specified further. Other generic sequences are shown in |
| CC | Q70466-68. Other specific peptides generated by these generic sequences |
| CC | are shown in R65151-54. TSARs are concatenated heterofunctional proteins |
| CC | or peptides, comprising at least two functional regions - a binding |
| CC | domain with affinity for a ligand and a second effector peptide portion |
| CC | that is chemically or biologically active. They may further comprise a |
| CC | linker peptide between the 2 domains. The oligonucleotides are also |
| CC | designed so that the expressed peptide contains 2 or 4 cysteine residues |
| CC | positioned in, or flanking, the unpredicted or variant residues. These |
| CC | residues confer some degree of conformational rigidity to the peptides. |
| CC | The TSARs or compns. comprising a TSAR binding domain can be used in |
| CC | vivo to deliver a chemically or biologically active moiety, eg. metal |
| CC | ion, radioisotope, peptide, toxin or enzyme, to the specific target or |
| CC | on the cell. They can also replace the function of macromolecules, eg. |
| CC | monoclonal or polyclonal antibodies and therefore circumvent the need for |
| CC | complex methods of hybridoma formation or in vivo antibody production. |
| CC | The TSARs are easily characterised and have designed activity allowing |
| CC | direct and rapid detection in a screening process. |
| CC | Sequence 114 BP; 0 A; 4 C; 4 G; 4 T; |
| QY | |

Query Match 2.1%; Score 40; DB 12; Length 114;
Best Local Similarity 8.9%; Pred. No. 1.32e-07;
Matches 10; Conservative 32; Mismatches 70; Indels 0; Gaps 0;

PI Fowlkes DM, Kay BK;
 DR WPI: 94-279739/34.
 PT screening a recombinant vector library expressing fusion proteins
 PT comprising a binding domain and an effector domain
 PS Disclosure; Page 35; 255pp; English.
 CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be
 CC represented as follows: X(NNB)6(TGC)(NNB)14(TGC)(NNB)3Y. X
 CC and Y are flanking restriction sites (X is not the same as Y) that are
 CC not specified further. Other generic sequences are shown in Q70466-68.
 CC Other specific peptides generated by these generic sequences are shown in
 CC Q65151-54. TSARs are concatenated heterofunctional proteins or peptides,
 CC comprising at least two functional regions - a binding domain with
 CC affinity for a ligand and a second effector peptide portion that is
 CC chemically or biologically active. They may further comprise a linker
 CC peptide between the 2 domains. The oligonucleotides are also designed so
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned
 CC in, or flanking, the unpredicted or variant residues. These residues
 CC confer some degree of conformational rigidity to the peptides. The TSARs
 CC or comsps. comprising a TSAR binding domain can be used in vivo to
 CC deliver a chemically or biologically active moiety, eg. metal ion,
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
 CC cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need
 CC for complex methods of hybridoma formation or in vivo antibody
 CC production. The TSARs are easily characterised and have designed
 CC activity allowing direct and rapid detection in a screening process.
 CC Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;
 Query Match 1.9%; Score 36; DB 12; Length 114;
 Best Local Similarity 4.5%; Pred. No. 1.31e-05;
 Matches 5; Conservative 33; Mismatches 74; Indels 0; Gaps 0;
 Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 52
 QY 69 GCGCGCGCTCCCGCTCGGAGCGCTGCTGCTGCGACGCCACTACCGCGAGGT 128
 Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 114
 QY 129 GCTGCGCGTGGCCAGCTTCGTGCGGCGCTGGGGCCCCAGGCGTGGCGGTG 180
 RESULT 15
 ID Q70469 standard; DNA; 114 BP.
 AC Q70469;
 DT 07-APR-1995 (first entry)
 DE Generic DNA sequence to generate a random TSAR peptide library.
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; generic; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 55...60
 FT /*tag= a
 FT /note= "this sequence represents 'z'; z can be a
 FT sequence of 6,9 or 12 nucleotides (see
 FT comments)"
 PN W09418318-A.
 PD 18-AUG-1994.
 PF 01-FEB-1994; U00977.
 PR 01-FEB-1993; US-013416.
 PR 30-DEC-1993; US-176500.
 PR 31-JAN-1994; US-189331.
 PA (UNNC-) UNIV NORTH CAROLINA.
 PI Fowlkes DM, Kay BK;
 DR WPI: 94-279739/34.
 PT Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PT comprising a binding domain and an effector domain
 PS Disclosure; Page 35; 255pp; English.
 CC Q70469 is a generic DNA sequence used to generate random TSAR peptide

CC This generic formula can be represented as follows: X(TGC)(NNB)10-
 CC (TGC)(NNB)62(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
 CC sites (X is not the same as Y) that are not specified further. This
 CC sequence generates peptides that are cloverleaf in structure. Other
 CC generic sequences are shown in Q70465-68. Other specific peptides
 CC generated by these generic sequences are shown in Q65150-54. TSARs are
 CC concatenated heterofunctional proteins or peptides, comprising at least
 CC two functional regions - a binding domain with affinity for a ligand and
 CC a second effector peptide portion that is chemically or biologically
 CC active. They may further comprise a linker peptide between the 2 domains.
 CC The oligonucleotides are also designed so that the expressed peptide
 CC contains 2 or 4 cysteine residues positioned in, or flanking, the
 CC unpredicted or variant residues. These residues confer some degree of
 CC conformational rigidity to the peptides. The TSARs or comsps. comprising
 CC a TSAR binding domain can be used in vivo to deliver a chemically or
 CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
 CC or enzyme, to the specific target or on the cell. They can also replace
 CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
 CC and therefore circumvent the need for complex methods of hybridoma
 CC formation or in vivo antibody production. The TSARs are easily
 CC characterised and have designed activity allowing direct and rapid
 CC detection in a screening process.
 CC Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;
 SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;
 Query Match 1.9%; Score 36; DB 12; Length 114;
 Best Local Similarity 7.9%; Pred. No. 1.31e-05;
 Matches 9; Conservative 31; Mismatches 74; Indels 0; Gaps 0;
 Db 1 tgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 60
 Cp 1085 TGCYCTTGTGCTGCTGAGGAGTAGAGGAGTGTGTCGCGGTACACCGGGGACAA 1026
 Db 61 nbnnbgtcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbgtc 114
 Cp 1025 GGCGTGTCCCGAGGACGTGTGTGGCGCGATGTGATGGGGGGCCCGCGTGTGTC 972
 Search completed: Thu Dec 24 15:06:48 1998
 Job time : 666 secs.

THIS PAGE BLANK (USPTO)

W P S R L H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 15:10:17 1998; MasPar time 3273.33 Seconds
Tabular output not generated. 628.529 Million cell updates/sec

Title: >US-08-951-733-13
Description: (1-1920) from US08951733.seq (1 of 2)
Perfect Score: 1920
N.A. Sequence: 1 CAGCGTCCGGCAGCGCTG.....CCAGGCCGCCCTGCTGACG 1920
Comp: GTGCGCAGGCCGCTCGGAC.....GGTCCGGCGGCGAGCTGC

Scoring table: TABLE default
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 1665728 seqs, 535777091 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-pending
1:P9 2:U6000 3:U6001 4:U6002 5:U6003 6:U6004 7:U6005
8:U6006 9:U6007 10:U6008 11:U7 12:U80 13:U81 14:U82
15:U83 16:U84A 17:U84B 18:U85 19:U86 20:U87 21:U88
22:U89 23:U90A 24:U90B 25:U91 26:NEWP 27:NEWU6 28:NEWU8
29:NEWU9

Statistics: Mean 10.626; Variance 4.576; scale 2.322

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|-----------------------------------|-----------|
| 1 | 1920 | 100.0 | 2848 | 22 | US-08-951- Sequence 13, Applicati | 0.00e+00 |
| 2 | 1920 | 100.0 | 3798 | 22 | US-08-951- Sequence 19, Applicati | 0.00e+00 |
| 3 | 1912 | 99.6 | 7029 | 22 | US-08-911- Sequence 1, Applicati | 0.00e+00 |
| 4 | 1909 | 99.4 | 3855 | 22 | US-08-912- Sequence 4, Applicati | 0.00e+00 |
| 5 | 1909 | 99.4 | 3855 | 22 | US-08-911- Sequence 18, Applicati | 0.00e+00 |
| 6 | 1909 | 99.4 | 4015 | 22 | US-08-913- Sequence 1, Applicati | 0.00e+00 |
| 7 | 1909 | 99.4 | 4015 | 24 | US-09-052- Sequence 1, Applicati | 0.00e+00 |
| 8 | 1909 | 99.4 | 4015 | 21 | US-08-854- Sequence 224, Applicat | 0.00e+00 |
| 9 | 1906 | 99.3 | 4023 | 23 | US-09-026- Sequence 35, Applicati | 0.00e+00 |
| 10 | 1596 | 83.1 | 4029 | 21 | US-08-854- Sequence 173, Applicat | 0.00e+00 |
| 11 | 1596 | 83.1 | 4029 | 21 | US-08-851- Sequence 173, Applicat | 0.00e+00 |
| 12 | 1497 | 78.0 | 4200 | 22 | US-08-913- Sequence 6, Applicati | 0.00e+00 |
| 13 | 453 | 23.6 | 3496 | 23 | US-09-043- Sequence 1, Applicati | 0.00e+00 |
| 14 | 281 | 14.6 | 3346 | 23 | US-09-026- Sequence 36, Applicati | 2.85e-234 |
| 15 | 259 | 13.5 | 535 | 22 | US-08-911- Sequence 52, Applicati | 6.37e-213 |

| | | | | | | | |
|----|-----|------|------|----|------------|------------------------|-----------|
| 16 | 245 | 12.8 | 550 | 22 | US-08-911- | Sequence 54, Applicati | 2.25e-199 |
| 17 | 232 | 12.0 | 2176 | 22 | US-08-912- | Sequence 3, Applicati | 7.99e-187 |
| 18 | 231 | 12.0 | 2171 | 21 | US-08-846- | Sequence 100, Applicat | 7.36e-186 |
| 19 | 231 | 12.0 | 2171 | 21 | US-08-851- | Sequence 100, Applicat | 7.36e-186 |
| 20 | 231 | 12.0 | 2171 | 21 | US-08-854- | Sequence 100, Applicat | 7.36e-186 |
| 21 | 227 | 11.8 | 389 | 21 | US-08-844- | Sequence 62, Applicati | 5.26e-182 |
| 22 | 227 | 11.8 | 389 | 21 | US-08-846- | Sequence 62, Applicati | 5.26e-182 |
| 23 | 227 | 11.8 | 389 | 21 | US-08-851- | Sequence 62, Applicati | 5.26e-182 |
| 24 | 227 | 11.8 | 389 | 22 | US-08-911- | Sequence 17, Applicati | 5.26e-182 |
| 25 | 227 | 11.8 | 389 | 21 | US-08-854- | Sequence 62, Applicati | 5.26e-182 |
| 26 | 227 | 11.8 | 389 | 22 | US-08-912- | Sequence 8, Applicatio | 5.26e-182 |
| 27 | 120 | 6.2 | 2651 | 23 | US-08-042- | Sequence 5, Applicatio | 1.23e-80 |
| 28 | 97 | 5.1 | 240 | 22 | US-08-912- | Sequence 7, Applicatio | 1.28e-59 |
| 29 | 69 | 3.6 | 477 | 11 | US-07-904- | Sequence 10, Applicati | 7.28e-35 |
| 30 | 69 | 3.6 | 477 | 11 | US-07-904- | Sequence 10, Applicati | 7.28e-35 |
| 31 | 69 | 3.6 | 1808 | 23 | US-09-042- | Sequence 4, Applicatio | 5.26e-35 |
| 32 | 68 | 3.5 | 578 | 18 | US-08-569- | Sequence 14, Applicati | 5.26e-34 |
| 33 | 68 | 3.5 | 578 | 19 | US-08-624- | Sequence 14, Applicati | 5.26e-34 |
| 34 | 67 | 3.5 | 7218 | 17 | US-08-466- | Sequence 14, Applicati | 3.79e-33 |
| 35 | 66 | 3.4 | 383 | 16 | US-08-446- | Sequence 8, Applicatio | 2.71e-32 |
| 36 | 66 | 3.4 | 383 | 16 | US-08-446- | Sequence 8, Applicatio | 2.71e-32 |
| 37 | 66 | 3.4 | 383 | 16 | US-08-446- | Sequence 8, Applicatio | 2.71e-32 |
| 38 | 66 | 3.4 | 383 | 15 | US-08-311- | Sequence 8, Applicatio | 2.71e-32 |
| 39 | 66 | 3.4 | 383 | 17 | US-08-451- | Sequence 8, Applicatio | 2.71e-32 |
| 40 | 66 | 3.4 | 383 | 16 | US-08-446- | Sequence 8, Applicatio | 2.71e-32 |
| 41 | 65 | 3.4 | 578 | 18 | US-08-569- | Sequence 14, Applicati | 1.93e-31 |
| 42 | 65 | 3.4 | 578 | 19 | US-08-624- | Sequence 14, Applicati | 1.93e-31 |
| 43 | 62 | 3.2 | 340 | 19 | US-07-617- | Sequence 9, Applicatio | 6.75e-29 |
| 44 | 62 | 3.2 | 477 | 11 | US-07-904- | Sequence 4, Applicatio | 6.75e-29 |
| 45 | 59 | 3.1 | 383 | 17 | US-08-451- | Sequence 8, Applicatio | 2.24e-26 |

ALIGNMENTS

RESULT 1
ID US-08-951-733-13 STANDARD; DNA; UNC; 2848 BP.
AC xxxxxx
DT
Sequence 13, Application US/08951733
Sequence 13, Application US/08951733
GENERAL INFORMATION:
CC APPLICANT: Harrington, Lea A.
CC APPLICANT: Robinson, Murray O.
CC TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS
CC NUMBER OF SEQUENCES: 44
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Angen Inc.
CC STREET: One Angen Center Drive
CC CITY: Thousand Oaks
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 91320-1789
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/951.733
CC FILING DATE: 16-OCT-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/873,039
CC FILING DATE: 11-JUN-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/751,189
CC FILING DATE: 15-NOV-1996
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Oleski, Nancy A.
CC REGISTRATION NUMBER: 34,688
CC REFERENCE/DOCKET NUMBER: A-433B
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (805) 447-6504

CC TELEFAX: (805) 499-8011
CC INFORMATION FOR SEQ ID NO: 13:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2848 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
SQ SEQUENCE 2848 BP; 437 A; 978 C; 945 G; 488 T; 0 OTHER.

Query Match 100.0%; Score 1920; DB 22; Length 2848;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 CACGGCTCCGGGACGCTGCTGCTGCGCAGCTGGAAGCCCTGGCCCCGGCCACC 60
Qy 1 CACGGCTCCGGGACGCTGCTGCTGCGCAGCTGGAAGCCCTGGCCCCGGCCACC 60

Db 61 CCCGGATGCGCGGCTCCCGCTGCGGAGCGCTGCGCTCCCTGCTGCGAGCCACTAC 120
Qy 61 CCCGGATGCGCGGCTCCCGCTGCGGAGCGCTGCGCTCCCTGCTGCGAGCCACTAC 120

Db 121 CGCAGGTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 180
Qy 121 CGCAGGTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 180

Db 181 GTGAGCGCGGGGACCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCTGCTGCTG 240
Qy 181 GTGAGCGCGGGGACCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCTGCTGCTG 240

Db 241 CCTGGGACGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Qy 241 CCTGGGACGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300

Db 301 GAGCTGTGGCCCGAGTCTGCGAGAGGTGTCGAGCGCGCGCGCGCGCGCGCGCG 360
Qy 301 GAGCTGTGGCCCGAGTCTGCGAGAGGTGTCGAGCGCGCGCGCGCGCGCGCGCG 360

Db 361 TTCGCTTCGCGCTGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Qy 361 TTCGCTTCGCGCTGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420

Db 421 GTGCGCACTTACCTGCGCAACAGGTGACGACGACGTGCGGCGCGCGCGCGCGCG 480
Qy 421 GTGCGCACTTACCTGCGCAACAGGTGACGACGACGTGCGGCGCGCGCGCGCGCG 480

Db 481 CTGCTGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 540
Qy 481 CTGCTGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 540

Db 541 TTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy 541 TTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

Db 601 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 601 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

Db 661 TGCAGCGCGCTGCAACATAGCTGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Qy 661 TGCAGCGCGCTGCAACATAGCTGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 720

Db 721 CCGGTGCGAGGAGCG 780
Qy 721 CCGGTGCGAGGAGCG 780

Db 781 AGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 781 AGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

Db 841 CCGGCGAGCGCTGCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy 841 CCGGCGAGCGCTGCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

Db 901 GCCGAGAGAGCCACCTCTTTGGAGGGTGCCTCTCTGGCAGCGGCCACTCCACCCATCC 960
Qy 901 GCCGAGAGAGCCACCTCTTTGGAGGGTGCCTCTCTGGCAGCGGCCACTCCACCCATCC 960

Db 961 GTGGGCGCGCAGCAGCAGCGGGGCCCCCATCCACATCGCGGCCACACGCTCCCTGGAC 1020
Qy 961 GTGGGCGCGCAGCAGCAGCGGGGCCCCCATCCACATCGCGGCCACACGCTCCCTGGAC 1020

Db 1021 ACGCCTTGTCCCCGGGTATACGCGGAGACCAAGCACTTCCTCTACTCTCCTCAGGCGACAAG 1080
Qy 1021 ACGCCTTGTCCCCGGGTATACGCGGAGACCAAGCACTTCCTCTACTCTCCTCAGGCGACAAG 1080

Db 1081 GAGCAGTGTGGGCGCTCCTTCTTACTCAGTCTCTGAGGCGCCAGCCTGACTGGCGCTCGG 1140
Qy 1081 GAGCAGTGTGGGCGCTCCTTCTTACTCAGTCTCTGAGGCGCCAGCCTGACTGGCGCTCGG 1140

Db 1141 AGGCTCTGTGAGACCATCTTTCTGGGTTCCAGGCGCTGGATGCCAGGAGCTCCCGCGAG 1200
Qy 1141 AGGCTCTGTGAGACCATCTTTCTGGGTTCCAGGCGCTGGATGCCAGGAGCTCCCGCGAG 1200

Db 1201 TTGCCCCGCTGCCCCAGCGCTACTGCGAAATGCGGCCCTGTTTCTGGAGCTGCTTGGG 1260
Qy 1201 TTGCCCCGCTGCCCCAGCGCTACTGCGAAATGCGGCCCTGTTTCTGGAGCTGCTTGGG 1260

Db 1261 AACACGCGCAGTGCCTTACGGGGTCTCCTCAAGACGCACTGCCCGCTGCGAGCTGCG 1320
Qy 1261 AACACGCGCAGTGCCTTACGGGGTCTCCTCAAGACGCACTGCCCGCTGCGAGCTGCG 1320

Db 1321 GTACACCCAGGAGCGGTGCTGTGCCCCGGGAGAGCCCGAGGCTCTGTGGGCGGCCCC 1380
Qy 1321 GTACACCCAGGAGCGGTGCTGTGCCCCGGGAGAGCCCGAGGCTCTGTGGGCGGCCCC 1380

Db 1381 GAGGAGGAGACACAGACCCCGCTGCTGCTGAGTGTGCTGCCAGCAGCAGCAGCCCC 1440
Qy 1381 GAGGAGGAGACACAGACCCCGCTGCTGCTGAGTGTGCTGCCAGCAGCAGCAGCCCC 1440

Db 1441 TGGCAGGTGTACGGCTTCTGTCGGGCGCTGCTGCGCGGCTGCTGCGCGGCTGCTGCG 1500
Qy 1441 TGGCAGGTGTACGGCTTCTGTCGGGCGCTGCTGCGCGGCTGCTGCGCGGCTGCTGCG 1500

Db 1501 GGCTCCAGGACACAGACCGCGCTTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGG 1560
Qy 1501 GGCTCCAGGACACAGACCGCGCTTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGG 1560

Db 1561 AAGCATGCCAAGCTCTCGCTGCGAGGCTGACGTGGAAGATGAGCTGCGGAGCTGCGCT 1620
Qy 1561 AAGCATGCCAAGCTCTCGCTGCGAGGCTGACGTGGAAGATGAGCTGCGGAGCTGCGCT 1620

Db 1621 TGGCTGCGCAGGAGCGCGGGGTTGGCTGTGTTCCGGCGCGAGACCGCTCTCGCTGAG 1680
Qy 1621 TGGCTGCGCAGGAGCGCGGGGTTGGCTGTGTTCCGGCGCGAGACCGCTCTCGCTGAG 1680

Db 1681 GAGATCTGCGCAAGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Qy 1681 GAGATCTGCGCAAGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740

Db 1741 TCTTTCTTTTATGTCGCGAGACCACTTTCAAAGAACAGGCTCTTTTCTACCGGAAG 1800
Qy 1741 TCTTTCTTTTATGTCGCGAGACCACTTTCAAAGAACAGGCTCTTTTCTACCGGAAG 1800

Db 1801 AGTCTGTGGAGCAAGTTGCAAGCAATTGGAATCAGACAGCACTTTGAAGAGGTTGAGCTG 1860
Qy 1801 AGTCTGTGGAGCAAGTTGCAAGCAATTGGAATCAGACAGCACTTTGAAGAGGTTGAGCTG 1860

Db 1861 CCGGAGCTGTGCGAAGCAGAGGTCAGGAGCATCGGAAAGCAGGCGCGCGCTGCTGAGC 1920
Qy 1861 CCGGAGCTGTGCGAAGCAGAGGTCAGGAGCATCGGAAAGCAGGCGCGCGCTGCTGAGC 1920

RESULT 2
ID US-08-951-733-19 STANDARD; DNA; UNC; 3798 BP.
AC xxxxxx

Sequence 19, Application US/08951733
Sequence 19, Application US/08951733
GENERAL INFORMATION:
APPLICANT: Harrington, Lea A.
APPLICANT: Robinson, Murray O.
TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: One Amgen Center Drive
CITY: Thousand Oaks
STATE: CA
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/951.733
FILING DATE: 16-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/873,039
FILING DATE: 11-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/751,189
FILING DATE: 15-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy A.
REGISTRATION NUMBER: 34,688
REFERENCE/DOCKET NUMBER: A-433B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (805) 447-6504
TELEFAX: (805) 499-8011
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 3798 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE 3798 BP: 613 A; 1310 C; 1213 G; 662 T; 0 OTHER.

Query Match 100.0%; Score 1920; DB 22; Length 3798;
Best Local Similarity 100.0%; Pred No. 0.00e+00;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 CACGCGTCCGGGACGCGTGGTCTGTCGCGACGTGGGAAGCCCTGGCCCCGGCCACC 61
Qy 1 CACGCGTCCGGGACGCGTGGTCTGTCGCGACGTGGGAAGCCCTGGCCCCGGCCACC 60
Db 62 CCGGGATGCGGGGCTCCCGCTGCGGAGCGTGGCGCTCCCTGCTGCGAGCCACTAC 121
Qy 61 CCGGGATGCGGGGCTCCCGCTGCGGAGCGCTGGCTCCCTGCTGCGAGCCACTAC 120
Db 122 CGGAGGTGCTGCGGCTGGCAGCTTCTGCGGCGCTGGGCGCCCGAGGCTGGCGCTG 181
Qy 121 CGGAGGTGCTGCGGCTGGCAGCTTCTGCGGCGCTGGGCGCCCGAGGCTGGCGCTG 180
Db 182 GTGACGCGGGGACCGCGGCTTCCGCGCGTGGTGGCCAGTGGCTGTGGCTG 241
Qy 181 GTGACGCGGGGACCGCGGCTTCCGCGCGTGGTGGCCAGTGGCTGTGGCTG 240
Db 242 CCCTGGGACGACGGCGCCCGCCCGCCCTTCCGCGAGGTGCTGCTGCTGAAG 301
Qy 241 CCCTGGGACGACGGCGCCCGCCCGCCCTTCCGCGAGGTGCTGCTGCTGAAG 300
Db 302 GAGCTGGTGGGCGGAGTCTGACAGGCTGTGCGAGCGCGCGGAGAAAGTGTGCTGCC 361
Qy 301 GAGCTGGTGGGCGGAGTCTGACAGGCTGTGCGAGCGCGCGGAGAAAGTGTGCTGCC 360

Db 362 TTGGCTTTCGCGTGTGGAGCGGGGCCCCGGGGCCCCCCCCGAGGCTTTCACACAGC 421
Qy 361 TTGGCTTTCGCGTGTGGAGCGGGGCCCCGGGGCCCCCCCCGAGGCTTTCACACAGC 420
Db 422 GTGGCAGCTACCTGCCCAACACGGTACGACGACTCGGGGAGCGGGGCTGGGG 481
Qy 421 GTGGCAGCTACCTGCCCAACACGGTACGACGACTCGGGGAGCGGGGCTGGGG 480
Db 482 CTGCTGCTGCGCGCTGGGCGAGCTGCTGTTTACCTGTGGCAGCTGCGCGCTC 541
Qy 481 CTGCTGCTGCGCGCTGGGCGAGCTGCTGTTTACCTGTGGCAGCTGCGCGCTC 540
Db 542 TTTGTGCTGTGCTCCAGCTGCGCTACAGTGTGGGGGCGCGCTGTACCACTC 601
Qy 541 TTTGTGCTGTGCTCCAGCTGCGCTACAGTGTGGGGGCGCGCTGTACCACTC 600
Db 602 GGCCTGCCACTAGGCGCGCGCCCGCCACACGCTAGTGGAGCGCTGTGGGA 661
Qy 601 GGCCTGCCACTAGGCGCGCGCCCGCCACACGCTAGTGGAGCGCTGTGGGA 660
Db 662 TGCAGACGGGCTTGGAAACCATAGCTCAGGAGCGCGGGTTCCTTGGGCTGCCAGCC 721
Qy 661 TGCAGACGGGCTTGGAAACCATAGCTCAGGAGCGCGGGTTCCTTGGGCTGCCAGCC 720
Db 722 CCGGGTGCAGAGGCGCGGGGCGAGTCCAGCGGAAGTCTGCGTTCGCCAAGAGGCC 781
Qy 721 CCGGGTGCAGAGGCGCGGGGCGAGTCCAGCGGAAGTCTGCGTTCGCCAAGAGGCC 780
Db 782 AGCGTGGCGTCCCTGAGCGGAGCGGAGCGGCGGCTTGGGAGGGGTCTTGGGCGCCAC 841
Qy 781 AGCGTGGCGTCCCTGAGCGGAGCGGAGCGGCGGCTTGGGAGGGGTCTTGGGCGCCAC 840
Db 842 CCGGGCAGGAGCGGTGGACGAGTACCGTGTGTGGTGTACCTGCCAGACCC 901
Qy 841 CCGGGCAGGAGCGGTGGACGAGTACCGTGTGTGGTGTACCTGCCAGACCC 900
Db 902 GCCGAAGAAGCCACTCTTTGGAGGGTGCCTCTCTGGCAGCGGCCACTCCACCCATCC 961
Qy 901 GCCGAAGAAGCCACTCTTTGGAGGGTGCCTCTCTGGCAGCGGCCACTCCACCCATCC 960
Db 962 GTGGGCGCCAGCAGCAGCGGGGCCCCCATTCACATCGGGGCGCCACGCTCCCTGGGAC 1021
Qy 961 GTGGGCGCCAGCAGCAGCGGGGCCCCCATTCACATCGGGGCGCCACGCTCCCTGGGAC 1020
Db 1022 ACGCCTTGTCCCGGTGTACGCGGAGACCAAGCAGCTTCTACTCTCAGGCGACAAG 1081
Qy 1021 ACGCCTTGTCCCGGTGTACGCGGAGACCAAGCAGCTTCTACTCTCAGGCGACAAG 1080
Db 1082 GAGCAGTGCAGGCTCTTCTACTCAGCTCTCTGAGGCGCCAGCCTGACTGGCGCTCGG 1141
Qy 1081 GAGCAGTGCAGGCTCTTCTACTCAGCTCTCTGAGGCGCCAGCCTGACTGGCGCTCGG 1140
Db 1142 AGGCTGTGAGAGCACTTTCTGGGTTTCAGGCGCTGGATGCCAGGACTTCCCGGAGG 1201
Qy 1141 AGGCTGTGAGAGCACTTTCTGGGTTTCAGGCGCTGGATGCCAGGAGTCTCCCGGAGG 1200
Db 1202 TTGCCCCGCTGCGCCAGCGCTACTGGCAATGGGCGCTGTTTCTGAGAGCTCTTTGGG 1261
Qy 1201 TTGCCCCGCTGCGCCAGCGCTACTGGCAATGGGCGCTGTTTCTGAGAGCTCTTTGGG 1260
Db 1262 AACACACGCGAGTGCCTTACGGGGTCTCTCTAAGACGCACTGCCCGCTGCGAGCTGCG 1321
Qy 1261 AACACACGCGAGTGCCTTACGGGGTCTCTCTAAGACGCACTGCCCGCTGCGAGCTGCG 1320
Db 1322 GTACACCCAGAGCGCGTGTGTGCCCCGGGAGAACCCAGGCTCTGTGGCGGCGCCCC 1381
Qy 1321 GTACACCCAGAGCGCGTGTGTGCCCCGGGAGAACCCAGGCTCTGTGGCGGCGCCCC 1380
Db 1382 GAGGAGGAGACAGACACCCCGTGGCTGTGCTGCTGCTGCTGCCGACGACAGAGCCCC 1441
Qy 1381 GAGGAGGAGACAGACACCCCGTGGCTGTGCTGCTGCTGCCGACGACAGAGCCCC 1440

Db 1442 TGGCAGGTGTACGGCTTCGTGGGGCCCTCGCTGCGCGGCTGGTGCCTCCAGGCTCTGG 1501
Qy 1441 TGGCAGGTGTACGGCTTCGTGGGGCCCTCGCTGCGCGGCTGGTGCCTCCAGGCTCTGG 1500
Db 1502 GCTCCAGGCAACAGACGCGCTTCCTCAGGAACACCAAGAGTTCATCTCCCTGGGG 1561
Qy 1501 GGCTCCAGGCAACAGACGCGCTTCCTCAGGAACACCAAGAGTTCATCTCCCTGGGG 1560
Db 1562 AAGCATGCCAAGCTCTCCTCAGAGAGTGTGAGTGTGGAAGAGTGTGCGGAGTGCCT 1621
Qy 1561 AAGCATGCCAAGCTCTCCTCAGAGAGTGTGAGTGTGGAAGAGTGTGCGGAGTGCCT 1620
Db 1622 TGGCTGCCAGGAGCCAGGGTTCGTGCTGCTGCGCGGAGTGTGAGTGTGAGTGTGAG 1681
Qy 1621 TGGCTGCCAGGAGCCAGGGTTCGTGCTGCTGCGCGGAGTGTGAGTGTGAGTGTGAG 1680
Db 1682 GAGATCTTGGCAAGTTCCTGACATGCTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAG 1741
Qy 1681 GAGATCTTGGCAAGTTCCTGACATGCTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAG 1740
Db 1742 TCTTCTTTTATGTCCAGGAGACACGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAG 1801
Qy 1741 TCTTCTTTTATGTCCAGGAGACACGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAG 1800
Db 1802 AGTGTCTGGAGCAAGTTCAGAGCAATTGGAATCAGACAGCACTTGAAGAGGTCAGCTG 1861
Qy 1801 AGTGTCTGGAGCAAGTTCAGAGCAATTGGAATCAGACAGCACTTGAAGAGGTCAGCTG 1860
Db 1862 CGGAGCTGTGCGAAGCAGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAG 1921
Qy 1861 CGGAGCTGTGCGAAGCAGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAG 1920

RESULT 3

ID US-08-911-312-1 STANDARD; DNA; UNC; 7029 BP.

AC xxxxxx

DT Sequence 1, Application US/08911312

DE Sequence 1, Application US/08911312

CC GENERAL INFORMATION:

CC APPLICANT: Cech, Thomas R.

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin B.

CC APPLICANT: Andrews, William

CC TITLE OF INVENTION: Telomerase Reverse Transcriptase

CC NUMBER OF SEQUENCES: 170

CC CORRESPONDENCE ADDRESSES: 170

CC ADDRESS: Townsend and Townsend and Crew LLP

CC STREET: Two Embarcadero Center, Eighth Floor

CC CITY: San Francisco

CC STATE: California

CC COUNTRY: USA

CC ZIP: 94111-3834

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/911,312

CC FILING DATE: 14-AUG-1997

CC CLASSIFICATION: 536

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/724,643

CC FILING DATE: 01-OCT-1996

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/844,419

CC FILING DATE: 18-APR-1997

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/846,017

CC

CC

CC

CC

CC

CC

CC

CC

CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Einhorn, Gregory P.
CC REGISTRATION NUMBER: 38,440
CC REFERENCE/DOCKET NUMBER: 015389-002500US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7029 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 782..4177
CC OTHER INFORMATION: /product= "human telomerase reverse
CC OTHER INFORMATION: transcriptase (hTRT)"
CC OTHER INFORMATION: /note= "cDNA contained in plasmid
CC OTHER INFORMATION: pGRN121"
SQ SEQUENCE 7029 BP; 1416 A; 2122 C; 2051 G; 1440 T; 0 OTHER.

Query Match

Best Local Similarity 99.8%; Score 1912; DB 22; Length 7029;

Matches 1915; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 718 CGCGTCGACGACGCGCTGCGTCCGTCGCGACGTGGAGAGCCCTGGCCCGCCACCCC 777
Qy 3 CGCGTCGCGGACGCGCTGCGTCCGTCGCGACGTGGAGAGCCCTGGCCCGCCACCCC 62
Db 778 CGCGATGCGCGCGCTCCCGCTCCCGACGCGTCCGTCCTGTCGCGACGACCTACCG 837
Qy 63 CGCGATGCGCGCGCTCCCGCTCCCGACGCGTCCGTCCTGTCGCGACGACCTACCG 122
Db 838 CGAGGTGCTGCGCTGCGCGACGCTTCGTCGCGCGCTGGGGCCCGAGGCTGGCGCTGGT 897
Qy 123 CGAGGTGCTGCGCTGCGCGACGCTTCGTCGCGCGCTGGGGCCCGAGGCTGGCGCTGGT 182
Db 898 GCAGCGGGGACCGCGGCTTCGCGCGCTGGTGGCCAGTGGCTGGTGGCTGGC 957
Qy 183 GCAGCGGGGACCGCGGCTTCGCGCGCTGGTGGCCAGTGGCTGGTGGCTGGC 242
Db 958 CTGGGACGACGCGCGCCCGCCCGCTCCCTCCGCGAGGTGCTGCTCCCTGAAGGA 1017
Qy 243 CTGGGACGACGCGCGCCCGCCCGCTCCCTCCGCGAGGTGCTGCTCCCTGAAGGA 302
Db 1018 GCTGGTGGCGCGAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGAGAACGCTGTGGCCTT 1077
Qy 303 GCTGGTGGCGCGAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGAGAACGCTGTGGCCTT 362
Db 1078 CGGCTTCGCGCTGCTGGACGGGCGCGGGGCGCGCGCGCGCGCGCTTCACCAACGCT 1137
Qy 363 CGGCTTCGCGCTGCTGGACGGGCGCGGGGCGCGCGCGCGCGCGCTTCACCAACGCT 422
Db 1138 GCGCAGCTACTGTCGCCAACACGCTGACCGACGCTGCGGGGAGCGGGGCTGGGGCT 1197
Qy 423 GCGCAGCTACTGTCGCCAACACGCTGACCGACGCTGCGGGGAGCGGGGCTGGGGCT 482
Db 1198 GCTGCTGGCGCGCTGGGCGACGAGTGTGCTGCTTCACCTGCTGGCAGCTGCTGCTT 1257

Qy 483 GCTGCTGCGCGCTGGCGAGACGCTGCTGGTTACCTGCTGGACACGCTGCGGCGCTTT 542
Db 1258 TGTGCTGTGTGCTCCAGCTCGGCTACCAAGTGTGCGGGCGCGCTGTACCAAGCTCGG 1317
Qy 543 TGTGCTGTGTGCTCCAGCTCGGCTACCAAGTGTGCGGGCGCGCTGTACCAAGCTCGG 602
Db 1318 CGTGTGCTGCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1377
Qy 603 CGTGTGCTGCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 662
Db 1378 CGAAGCGGCTTGAACACATAGCTGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1437
Qy 663 CGAAGCGGCTTGAACACATAGCTGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 722
Db 1438 GGTGTGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1497
Qy 723 GGTGTGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 782
Db 1498 GCTGTGCGCTGCGGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1557
Qy 783 GCTGTGCGCTGCGGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 842
Db 1558 GGCAGGAGCGCTGACGCGAGTGACCGGTGTTCTGTGTGTGTACCTGCGGCGGCGGCGG 1617
Qy 843 GGCAGGAGCGCTGACGCGAGTGACCGGTGTTCTGTGTGTGTACCTGCGGCGGCGGCGG 902
Db 1618 CGAAGAGCCACCTCTTTGAGGCTGCGCTCTCTGCGCAGCGGCGGCGGCGGCGGCGG 1677
Qy 903 CGAAGAGCCACCTCTTTGAGGCTGCGCTCTCTGCGCAGCGGCGGCGGCGGCGGCGG 962
Db 1678 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1737
Qy 963 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1022
Db 1738 GCCTGTGCTCCCGGTGACGCGGAGACCAAGCACTTCTTACCTCTCAGCGGCGGCGG 1797
Qy 1023 GCCTGTGCTCCCGGTGACGCGGAGACCAAGCACTTCTTACCTCTCAGCGGCGGCGG 1082
Db 1798 CGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1857
Qy 1083 CGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1142
Db 1858 GCTGTGTGAGACCATCTTTCTGGGTTCAGGCGCTGTGATGCGGAGGATCTCCCGGAGGTT 1917
Qy 1143 GCTGTGTGAGACCATCTTTCTGGGTTCAGGCGCTGTGATGCGGAGGATCTCCCGGAGGTT 1202
Db 1918 GCGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1977
Qy 1203 GCGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1262
Db 1978 CCACGCGAGTGTGCGGCGGCTGCTCTCAAGACGCACTGCGGCGTGTGAGGCTGCGGT 2037
Qy 1263 CCACGCGAGTGTGCGGCGGCTGCTCTCAAGACGCACTGCGGCGTGTGAGGCTGCGGT 1322
Db 2038 CACCGGCGAGCGGCTGTGTGCGGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2097
Qy 1323 CACCGGCGAGCGGCTGTGTGCGGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1382
Db 2098 GGAGGAGACACAGACCGGCTGCGCTGGTGTGCTGCGGCGGCGGCGGCGGCGGCGGCGG 2157
Qy 1383 GGAGGAGACACAGACCGGCTGCGCTGGTGTGCTGCGGCGGCGGCGGCGGCGGCGGCGG 1442
Db 2158 CGAGTGTGTGCGGCTGCGGCGGCTGCTGCGGCGGCTGCTGCGGCGGCGGCGGCGGCGG 2217
Qy 1443 CGAGTGTGTGCGGCTGCGGCGGCTGCTGCGGCGGCTGCTGCGGCGGCGGCGGCGGCGG 1502
Db 2218 CTCAGGCGACACAGACCGGCTTCTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGA 2277
Qy 1503 CTCAGGCGACACAGACCGGCTTCTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGA 1562
Db 2278 GCATCCGAGCTCTGCTGCGAGGAGCTGAGTGGAGAGTGTGCGGCGGAGTGTGCGGCTG 2337

Qy 1563 GCATCCCAAGCTCTGCTGCGAGGAGCTGAGCGTGGAAATGAGCGTGGGGGAGCTGCGGCTG 1622
Db 2338 GCTCGGAGGAGCCAGGGGTTGGCTGTGTCCGGCGCGAGAGCACCGCTGTGCGTGAGGA 2397
Qy 1623 GCTCGGAGGAGCCAGGGGTTGGCTGTGTCCGGCGCGAGAGCACCGCTGTGCGTGAGGA 1682
Db 2398 GATCCTGGCCAAAGTCTGCTGCACTGGCTGTGAGTGTGTACGTCGTGAGCTGCTCAGGTC 2457
Qy 1683 GATCCTGGCCAAAGTCTGCTGCACTGGCTGTGAGTGTGTACGTCGTGAGCTGCTCAGGTC 1742
Db 2458 TTTCTTTTATGTACGAGGACACGCTTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAG 2517
Qy 1743 TTTCTTTTATGTACGAGGACACGCTTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAG 1802
Db 2518 TGTCTGAGCAAGTTGCAAGCATTGGAATCAGACAGCACTTGAAGAGGTTGAGCTGCG 2577
Qy 1803 TGTCTGAGCAAGTTGCAAGCATTGGAATCAGACAGCACTTGAAGAGGTTGAGCTGCG 1862
Db 2578 GGAGCTGTGCGGAACAGAGGTGAGGAGCATCGGGAAGCCAGGCGGCGGCTGCTGAGC 2635
Qy 1863 GGAGCTGTGCGGAACAGAGGTGAGGAGCATCGGGAAGCCAGGCGGCGGCTGCTGAGC 1920

RESULT 4
ID US-08-912-951-4 STANDARD; DNA; UNC; 3855 BP.
AC xxxxxx
DT
Sequence 4, Application US/08912951
DE Sequence 4, Application US/08912951
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002600US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 3855 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: -
CC LOCATION: 1..3855 /note= "nucleic acid sequence with an
CC OTHER INFORMATION: /product= "delta-182 variant
CC OTHER INFORMATION: open reading frame encoding a delta-182
CC OTHER INFORMATION: variant polypeptide"
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 56..2479
CC OTHER INFORMATION:
CC OTHER INFORMATION: polypeptide"
SQ SEQUENCE 3855 BP; 651 A; 1300 C; 1226 G; 678 T; 0 OTHER.

Query Match 99.4%; Score 1909; DB 22; Length 3855;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GCAGCGCTGCGTCTGCGCACGTGGAGCCCTGCGCCCGCCAGCCACCCCGCGATGCC 60
Qy 12 GCAGCGCTGCGTCTGCGCACGTGGAGCCCTGCGCCCGCCAGCCACCCCGCGATGCC 71
Db 61 GCGCGCTCCCGCTCCGAGCGCTGCGTCTGCGCACGTGGAGCCCTGCGCCCGCGATGCC 120
Qy 72 GCGCGCTCCCGCTCCGAGCGCTGCGTCTGCGCACGTGGAGCCCTGCGCCCGCGATGCC 131
Db 121 GCGCGTGGCCAGCTTCTGCGCGCGCTGGGGCCCGCCAGCGCTGGCGTGGCGAGCGG 180
Qy 132 GCGCGTGGCCAGCTTCTGCGCGCGCTGGGGCCCGCCAGCGCTGGCGTGGCGAGCGG 191
Db 181 GGACCCGGCGGCTTCCGCGCGCTGGTGGCCAGTGGCTGGTGGCGCTGGCGAGCGC 240
Qy 192 GGACCCGGCGGCTTCCGCGCGCTGGTGGCCAGTGGCTGGTGGCGCTGGCGAGCGC 251
Db 241 ACGCGCGCGCGCGCGCGCGCTTCTCCGCGAGTGTCTGCGTGAAGAGCTGGTGGC 300
Qy 252 ACGCGCGCGCGCGCGCGCGCTTCTCCGCGAGTGTCTGCGTGAAGAGCTGGTGGC 311
Db 301 CCGAGTGTGAGAGCTGTGCGAGCGCGCGCGCGAGAGCTGTGCGCTTGGCGTTCGC 360
Qy 312 CCGAGTGTGAGAGCTGTGCGAGCGCGCGCGCGAGAGCTGTGCGCTTGGCGTTCGC 371
Db 361 GCTGTGGACGGGCGCGCGCGCGCGCGCTTCCAGAGCTTCCAGAGCTGGCGAGCTA 420
Qy 372 GCTGTGGACGGGCGCGCGCGCGCGCGCTTCCAGAGCTTCCAGAGCTGGCGAGCTA 431
Db 421 CCGTCCCAACACGGTGACCGAGCTGCGGGGAGCGGGGCGTGGGGCTGCTGCTGCG 480
Qy 432 CCGTCCCAACACGGTGACCGAGCTGCGGGGAGCGGGGCGTGGGGCTGCTGCTGCG 491
Db 481 CCGCGTGGCGAGAGCTGTGCTGCTGCGAGCTGCGGGCTTGTGCTGCTGCTGCTGCT 540
Qy 492 CCGCGTGGCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551
Db 541 GCTCCCGAGCTGCGCTTACAGTGTGGGGCGCGCGCTGTACAGCTGCGGGCTGCCAC 600
Qy 552 GCTCCCGAGCTGCGCTTACAGTGTGGGGCGCGCGCTGTACAGCTGCGGGCTGCCAC 611

Db 601 TCAGGCCGCGCGCGCGCGCGCGCTAGTGACCCCGGAGGCGTCTGGGATCGGAACGGGC 660
Qy 612 TCAGGCCGCGCGCGCGCGCGCGCTAGTGACCCCGGAGGCGTCTGGGATCGGAACGGGC 671
Db 661 CTGGAACCATAGCGTCAAGGAGCGCGGGTCCCTTGGGCTTGCAGCCCCGGTGGAG 720
Qy 672 CTGGAACCATAGCGTCAAGGAGCGCGGGTCCCTTGGGCTTGCAGCCCCGGTGGAG 731
Db 721 GAGCGCGGGGCGAGTGCAGCCGAACTCTGCCGTTGCCAAGAGGCCAGCGCTGGCGC 780
Qy 732 GAGCGCGGGGCGAGTGCAGCCGAACTCTGCCGTTGCCAAGAGGCCAGCGCTGGCGC 791
Db 781 TGCCCCCTGAGCCGAGCGGAGCGCCGTTGGGCAAGGCTTCTGGGCCACCCCGGCGAGAC 840
Qy 792 TGCCCCCTGAGCCGAGCGGAGCGCCGTTGGGCAAGGCTTCTGGGCCACCCCGGCGAGAC 851
Db 841 GCGTGGACCGAGTACCGTGGTCTGTGTGGTCTACCTGCCAGACCCCGGAGAGC 900
Qy 852 GCGTGGACCGAGTACCGTGGTCTGTGTGGTCTACCTGCCAGACCCCGGAGAGC 911
Db 901 CACCTCTTTGAGGCTGCGCTCTCTGCGACGCGCACTTCCACCATCCGTTGGCGCGCA 960
Qy 912 CACCTCTTTGAGGCTGCGCTCTCTGCGACGCGCACTTCCACCATCCGTTGGCGCGCA 971
Db 961 GCACACGCGGCG 1020
Qy 972 GCACACGCGGCG 1031
Db 1021 CCCGGTGTAGCGCGAGACCGAGCTTCTCTACTCTCTAGCGCGACAGGAGCGAGCTGCG 1080
Qy 1032 CCCGGTGTAGCGCGAGACCGAGCTTCTCTACTCTCTAGCGCGACAGGAGCGAGCTGCG 1091
Db 1081 GCCCTCTTCTACTCTAGCTCTCTGAGCGCCAGCTGACTGCGCTCGGAGGCTCGTGGGA 1140
Qy 1092 GCCCTCTTCTACTCTAGCTCTCTGAGCGCCAGCTGACTGCGCGCTCGGAGGCTCGTGGGA 1151
Db 1141 GACCATCTTCTGGGTTCCAGGCGCTGATGCCAGGACTTCCCGCGAGGTTGCCCGCGCT 1200
Qy 1152 GACCATCTTCTGGTTCAGGCGCTGATGCCAGGACTTCCCGCGAGGTTGCCCGCGCT 1211
Db 1201 GCGCGCGGCTACTTGGCAATGCGCGCGCTGTTCTTGGAGCTGCTTGGGAGCGAGCGCA 1260
Qy 1212 GCGCGCGGCTACTTGGCAATGCGCGCGCTGTTCTTGGAGCTGCTTGGGAGCGAGCGCA 1271
Db 1261 GTGCGCGCTACGGGCTGCTCTCAAGAGCGACTGCGCGCTGCGAGCTGCGGTCACCCAGC 1320
Qy 1272 GTGCGCGCTACGGGCTGCTCTCAAGAGCGACTGCGCGCTGCGAGCTGCGGTCACCCAGC 1331
Db 1321 AGCGGCTGTGTGTCGCGGGAGAGCGCGCGCTGTGTGCGCGCGCGCGCGCGCGCGAGGA 1380
Qy 1332 AGCGGCTGTGTGTCGCGGGAGAGCGCGCGCTGTGTGCGCGCGCGCGCGCGCGAGGA 1391
Db 1381 CACAGACCGCGCTGCGCTGCGCGCGCTGCTCCGCGAGCGAGCGCGCTTGGGAGTGTGA 1440
Qy 1392 CACAGACCGCGCTGCGCTGCGCGCGCTGCTCCGCGAGCGAGCGCGCTTGGGAGTGTGA 1451
Db 1441 CCGTGTGCGCGCGCTGCGCTGCGCGCGCTGCTCCGCGAGCGCTTGGGAGTGTGA 1500
Qy 1452 CCGTGTGCGCGCGCTGCGCTGCGCGCGCTGCTCCGCGAGCGCTTGGGAGTGTGA 1511
Db 1501 CAACGAACCGCGCTTCTCAGGAACACCAAGAACTTCTCTCCCTGGGGAAGCATGCCAA 1560
Qy 1512 CAACGAACCGCGCTTCTCAGGAACACCAAGAACTTCTCTCCCTGGGGAAGCATGCCAA 1571
Db 1561 GCTCTCGCTCAGAGCTGAGCTGGAAGATGAGCTGCGGGACTGCGCTTGGCTGCGGAG 1620
Qy 1572 GCTCTCGCTCAGAGCTGAGCTGGAAGATGAGCTGCGGGACTGCGCTTGGCTGCGGAG 1631
Db 1621 GAGCCAGGGGTTGGCTGTGTTCGCGCGCGAGACCGCTGTGCTGAGGAGATCTCTGCG 1680
Qy 1632 GAGCCAGGGGTTGGCTGTGTTCGCGCGCGAGACCGCTGTGCTGAGGAGATCTCTGCG 1691

| | | | |
|---|--------------------|--|-----|
| CC | NAME/KEY: | CDS | |
| CC | LOCATION: | 56..3454 | |
| CC | OTHER INFORMATION: | /product= "hTfT" | |
| CC | OTHER INFORMATION: | /note= "human telomerase reverse | |
| CC | OTHER INFORMATION: | transcriptase (hTfT) catalytic protein | |
| CC | OTHER INFORMATION: | component" | |
| SQ | SEQUENCE | 4015 BP; 663 A; 1363 G; 1275 C; 714 T; 0 OTHER.. | |
| Query Match 99.4%; Score 1909; DB 22; Length 4015; | | | |
| Best Local Similarity 100.0%; Pred. No. 0.00e+00; | | | |
| Matches 1909; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
| Db | 1 | GCACGCTGCTGCTGCTGCGACGTGGGAAGCCCTGGCCCGCCACCCCGCGATGCC | 60 |
| QY | 12 | GCACGCTGGCTGCTGCTGCGACGTGGGAAGCCCTGGCCCGCCACCCCGCGATGCC | 71 |
| Db | 61 | GCAGCTCCCGCTGCGGAGCGTGGCTCCCTGCTGCGAGCCACTACCGGAGGTGCT | 120 |
| QY | 72 | GCAGCTCCCGCTGCGGAGCGTGGCTCCCTGCTGCGAGCCACTACCGGAGGTGCT | 131 |
| Db | 121 | GCAGCTGGCAGCTGCTGGGCGCTGGGCGCCCGCCAGGCTGGGCGCTGGTGCAGCGCG | 180 |
| QY | 132 | GCAGCTGGCAGCTGCTGGGCGCTGGGCGCCCGCCAGGCTGGGCGCTGGTGCAGCGCG | 191 |
| Db | 181 | GGACCGCGGCTTTCCGCGGCTGGTGGCCACTGCTGGTGGCTGGCTGGGAGCG | 240 |
| QY | 192 | GGACCGCGGCTTTCCGCGGCTGGTGGCCACTGCTGGTGGCTGGCTGGGAGCG | 251 |
| Db | 241 | ACGCGCGCGCCCGCCCGCTCTTCCGCGAGGTGCTGCTGAAAGAGCTGGTGGC | 300 |
| QY | 252 | ACGCGCGCGCCCGCCCGCTCTTCCGCGAGGTGCTGCTGAAAGAGCTGGTGGC | 311 |
| Db | 301 | CCGAGTGTGAGAGGTGTGCGAGCGGCGCGGAGGAGTGTGCTGGCTTGGCTTGGC | 360 |
| QY | 312 | CCGAGTGTGAGAGGTGTGCGAGCGGCGCGGAGGAGTGTGCTGGCTTGGCTTGGC | 371 |
| Db | 361 | GCTCTGAGCGGCGCGCGGCGCGCCCGCGAGGCTTACACACCGCTGGCGAGCTA | 420 |
| QY | 372 | GCTCTGAGCGGCGCGCGGCGCGCCCGCGAGGCTTACACACCGCTGGCGAGCTA | 431 |
| Db | 421 | CCTGCCAACACGCTGACCGACACTGCGGGGAGGGGCGTGGGGCTGCTGCTGCG | 480 |
| QY | 432 | CCTGCCAACACGCTGACCGACACTGCGGGGAGGGGCGTGGGGCTGCTGCTGCG | 491 |
| Db | 481 | CCGCGTGGCGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 540 |
| QY | 492 | CCGCGTGGCGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 551 |
| Db | 541 | GGCTCCAGCTGCGCTACCAAGTGTGGGCGCGCGCTGTACCACTGCGCGCTGCCAC | 600 |
| QY | 552 | GGCTCCAGCTGCGCTACCAAGTGTGGGCGCGCGCTGTACCACTGCGCGCTGCCAC | 611 |
| Db | 601 | TCAGCGCGCGCGCGCGCGCTAGTGGACCCCGAGGCGCTGCGGATGCGAACCGGC | 660 |
| QY | 612 | TCAGCGCGCGCGCGCGCGCTAGTGGACCCCGAGGCGCTGCGGATGCGAACCGGC | 671 |
| Db | 661 | CTGAACCATAGCTACGAGGCGCGGGTCCCTTGGGCTGCGAGCCCGGCTGGCGAG | 720 |
| QY | 672 | CTGAACCATAGCTACGAGGCGCGGGTCCCTTGGGCTGCGAGCCCGGCTGGCGAG | 731 |
| Db | 721 | GAGCGCGGGGCGAGTGCACCGAGCTGCGGCTGCGGCGCGCGCTGCGGCTGGCGC | 780 |
| QY | 732 | GAGCGCGGGGCGAGTGCACCGAGCTGCGGCTGCGGCGCGCGCTGCGGCTGGCGC | 791 |
| Db | 781 | TGCCCTGAGCGCGGAGCGCGCGTGGGCGAGGGTTCCTGGGCGCCACCGCGGCGAGC | 840 |
| QY | 792 | TGCCCTGAGCGCGGAGCGCGCGTGGGCGAGGGTTCCTGGGCGCCACCGCGGCGAGC | 851 |
| Db | 841 | CGGTGGACCGAGTACCGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 900 |
| QY | 852 | CGGTGGACCGAGTACCGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 911 |
| Db | 901 | CACCTCTTTGGAGGGTGGCTCTCTGGCACCGCGCCACTCCACCCTCCCTGGGCGCGCA | 960 |

RESULT 7
ID US-09-052-919-1 STANDARD; DNA; UNC; 4015 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/09052919

| | | | |
|----|------|---|------|
| QY | 912 | CACCTCTTTGGAGGGTGGCTCTCTGGCAGCGCCACTCCACCCATCCGTGGGCGCGCA | 971 |
| Db | 961 | GCACCAAGCGGGCCCCCATCCACATCGGGGCCACACAGCTCCCTGGGACACGCTTGTCC | 1020 |
| QY | 972 | GCACCAAGCGGGCCCCCATCCACATCGGGGCCACACAGCTCCCTGGGACACGCTTGTCC | 1031 |
| Db | 1021 | CCCGGTGTAGCGCGAGACCAAGACATTCCTCTACTCTCAGGCGACAAGGAGAGCTGG | 1080 |
| QY | 1032 | CCCGGTGTAGCGCGAGACCAAGACATTCCTCTACTCTCAGGCGACAAGGAGAGCTGG | 1091 |
| Db | 1081 | GCCTCTCTCTACTCAGCTCTCTGAGGCCACGCTGACTGGCGCTCGGAGGCTCGTGA | 1140 |
| QY | 1092 | GCCTCTCTCTACTCAGCTCTCTGAGGCCACGCTGACTGGCGCTCGGAGGCTCGTGA | 1151 |
| Db | 1141 | GACCATCTTTCTGGGTTCAGGCCCTGGATGCCAGGAGCTCCCGCAGGTTGCCCGCCT | 1200 |
| QY | 1152 | GACCATCTTTCTGGGTTCAGGCCCTGGATGCCAGGAGCTCCCGCAGGTTGCCCGCCT | 1211 |
| Db | 1201 | GCCCCAGCGTACTGGCAATGCGGCCCTGTTTCTGGAGCTGCTTGGGAACACAGCGCA | 1260 |
| QY | 1212 | GCCCCAGCGTACTGGCAATGCGGCCCTGTTTCTGGAGCTGCTTGGGAACACAGCGCA | 1271 |
| Db | 1261 | GTCCCCCTAGCGGCTGCTCCTCAAGACGCACTGCCCGCTCGAGCTGCGGTCAACCGAGC | 1331 |
| QY | 1272 | GTCCCCCTAGCGGCTGCTCCTCAAGACGCACTGCCCGCTCGAGCTGCGGTCAACCGAGC | 1331 |
| Db | 1321 | AGCGGCTGTCTGTCGCCGGGAGAGCCCAAGGCTCTGTGGCGGCCCGCCAGGAGGAGGA | 1380 |
| QY | 1332 | AGCGGCTGTCTGTCGCCGGGAGAGCCCAAGGCTCTGTGGCGGCCCGCCAGGAGGAGGA | 1391 |
| Db | 1381 | CACAGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 1440 |
| QY | 1392 | CACAGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 1451 |
| Db | 1441 | CGCTCTGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 1500 |
| QY | 1452 | CGCTCTGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 1511 |
| Db | 1501 | CAACGAACGCGCTTCTCCTCAGGAACACCAAGAGTTCATCTCCTCGGGAAGCATGCCAA | 1560 |
| QY | 1512 | CAACGAACGCGCTTCTCCTCAGGAACACCAAGAGTTCATCTCCTCGGGAAGCATGCCAA | 1571 |
| Db | 1561 | GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 1620 |
| QY | 1572 | GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 1631 |
| Db | 1621 | GAGCCAGGGGTTGGCTGCTTCCGGCCGAGACACCGCTCGCTGAGGAGATCCCTGGC | 1680 |
| QY | 1632 | GAGCCAGGGGTTGGCTGCTTCCGGCCGAGACACCGCTCGCTGAGGAGATCCCTGGC | 1691 |
| Db | 1681 | CAAGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 1740 |
| QY | 1692 | CAAGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 1751 |
| Db | 1741 | TGTCAGGAGACACGCTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAG | 1800 |
| QY | 1752 | TGTCAGGAGACACGCTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAG | 1811 |
| Db | 1801 | CAAGTTCTCAGAGCATTTGAATCAGACGACTTGAAGAGGGTGCAGCTGGGAGGCTGTC | 1860 |
| QY | 1812 | CAAGTTCTCAGAGCATTTGAATCAGACGACTTGAAGAGGGTGCAGCTGGGAGGCTGTC | 1871 |
| Db | 1861 | GGAAGCAGAGGCTCAGGAGCATCGGGAAGCAGCGCCCGCTGCTGACG | 1909 |
| QY | 1872 | GGAAGCAGAGGCTCAGGAGCATCGGGAAGCAGCGCCCGCTGCTGACG | 1920 |

Sequence 1, Application US/09052919
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Antisense Compositions for Detecting and
TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/052,919
FILING DATE: 31-MAR-1998
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/974,549
FILING DATE: 19-NOV-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Parent, Annette S.
REGISTRATION NUMBER: 42,058
REFERENCE/DOCKET NUMBER: 015389-003600US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

CC LENGTH: 4015 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 56..3454
CC OTHER INFORMATION: /product= "human telomerase reverse
CC OTHER INFORMATION: transcriptase (hTERT)"
SQ SEQUENCE 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 OTHER.

Query Match 99.4%; Score 1909; DB 24; Length 4015;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GCAGCGCTGGTCTGCTGCGCAGCTGGGAAGCCCTGGCCCGCCGACACCCCGCGATGCC 60
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
12 GCAGCGCTGGTCTGCTGCGCAGCTGGGAAGCCCTGGCCCGCCGACACCCCGCGATGCC 71
Db 61 GCGCGCTCCCGCTGCGGAGCCGCTCCCTGCTGCGCAGCAGCAGCTACCGCAGGTGCT 120
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
72 GCGCGCTCCCGCTGCGGAGCCGCTCCCTGCTGCGCAGCAGCAGCTACCGCAGGTGCT 131
Db 121 GCGCGTGGCCACGTTGCTGCGGCGCCTGGGCGCCCGAGGGCTGGCGGTGTGTGAGCGCG 180
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
132 GCGCGTGGCCACGTTGCTGCGGCGCCTGGGCGCCCGAGGGCTGGCGGTGTGTGAGCGCG 191
Db 181 GGACCGCGCGCTTCCGCGCGCTGGTGGCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
192 GGACCGCGCGCTTCCGCGCGCTGGTGGCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 251
Db 241 ACGCCCGCCCGCGCCCGCCCTCCCTCCCGCAGAGTGTCTGCTGAAGAGAGTGGTGGC 300
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
252 ACGCCCGCCCGCGCCCGCCCTCCCTCCCGCAGAGTGTCTGCTGAAGAGAGTGGTGGC 311
Db 301 CCGAGTGTGTGAGAGGCTGTGCGAGCGCGCGGCGGAAGACGTGTGCGCTTCGGCTTCGC 360
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
312 CCGAGTGTGTGAGAGGCTGTGCGAGCGCGCGGCGGAAGACGTGTGCGCTTCGGCTTCGC 371
Db 361 GCTGCTGGAGCGGCG 420
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
372 GCTGCTGGAGCGGCG 431
Db 421 CCTGCCCAACACAGCTGACACGACACTGCGCGGGAGCGGGGCGGTGGGGGTGTGCTGCG 480
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
432 CCTGCCCAACACAGCTGACACGACACTGCGCGGGAGCGGGGCGGTGGGGGTGTGCTGCG 491
Db 481 CCGGTGGGCGACAGCTGTGTTTCACTGTGTGGCAGCTGCGCGCTCTTTTGTGCTGGT 540
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
492 CCGGTGGGCGACAGCTGTGTTTCACTGTGTGGCAGCTGCGCGCTCTTTTGTGCTGGT 551
Db 541 GGCTCCAGCTGCGCTTACAGAGTGTGCGGCGCGCGCGCTTACAGCTCGCGCGCTGCCAC 600
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
552 GGCTCCAGCTGCGCTTACAGAGTGTGCGGCGCGCGCGCTTACAGCTCGCGCGCTGCCAC 611
Db 601 TCAGGCG 660
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
612 TCAGGCG 671
Db 661 CTGGAACCATAGCGTCAAGGAGCGCGGGTCCCGCTGGGCGCTGCCAGCCCGGGTGGAG 720
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
672 CTGGAACCATAGCGTCAAGGAGCGCGGGTCCCGCTGGGCGCTGCCAGCCCGGGTGGAG 731
Db 721 GAGCGCGGGGCGAGTCCAGCGAGCTGTGCGTTGCCCAAGAGCGCGAGGCGTGGCGC 780
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
732 GAGCGCGGGGCGAGTCCAGCGAGCTGTGCGTTGCCCAAGAGCGCGAGGCGTGGCGC 791
Db 781 TGCCCGCTGAGCGGAGCGCGCGCGCTTGGGCGAGGGGTCTTGGGCGCCACCGCGGAGAC 840
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
792 TGCCCGCTGAGCGGAGCGCGCGCGCTTGGGCGAGGGGTCTTGGGCGCCACCGCGGAGAC 851
Db 841 GCGTGGACCGAGTGACCGCTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900

|||||
QY 852 GCGTGGACGAGTACCGTGGTTCTGTGTGTGTACCTCCAGACCCGCGGAGAGC 911
Db 901 CACCTCTTTGAGGGTGGCTCTCTGGACGCGCCACTCCACCCATCCGTGGGCCGCCA 960
QY 912 CACCTCTTTGAGGGTGGCTCTCTGGACGCGCCACTCCACCCATCCGTGGGCCGCCA 971
Db 961 GCACACGCGGGCCGCCATCCATCCGCGCCACCGAGTCCCTGGACAGCCCTTGTC 1020
QY 972 GCACACGCGGGCCGCCATCCATCCGCGCCACCGAGTCCCTGGACAGCCCTTGTC 1031
Db 1021 CCCGGTGTACCCGAGACCAAGCACTTCTCTACTCTCTCAGCGCACCAAGGAGCAGCTGCG 1080
QY 1032 CCCGGTGTACCCGAGACCAAGCACTTCTCTACTCTCTCAGCGCACCAAGGAGCAGCTGCG 1091
Db 1081 GCCCTCTCTCTACTAGCTCTCTGAGGCCAGCCTGACTGGCCCTCGGAGGCTCGTGGA 1140
QY 1092 GCCCTCTCTCTACTAGCTCTCTGAGGCCAGCCTGACTGGCCCTCGGAGGCTCGTGGA 1151
Db 1141 GACCATCTTTCTGGTTCCAGGCCCTGGATGCCAGGAGTCCCGCGCAGGTTGCCCGCCCT 1200
QY 1152 GACCATCTTTCTGGTTCCAGGCCCTGGATGCCAGGAGTCCCGCGCAGGTTGCCCGCCCT 1211
Db 1201 GCCCAGCGCTACTGGCAATTCGCGCCCTCTGTCTGGAGCTGCTTGGGAACACACGCGCA 1260
QY 1212 GCCCAGCGCTACTGGCAATTCGCGCCCTCTGTCTGGAGCTGCTTGGGAACACACGCGCA 1271
Db 1261 GTGCCCTCTACGGGTGTCTCTCAAGAGCACTGCCCGCTGGAGCTGCGGTCAACCCCGAG 1320
QY 1272 GTGCCCTCTACGGGTGTCTCTCAAGAGCACTGCCCGCTGGAGCTGCGGTCAACCCCGAG 1331
Db 1321 AGCCGGTGTCTGTCGCGGGAGAGCCCGAGGCTCTGTGGCGGCCCGCGAGGAGGGA 1380
QY 1332 AGCCGGTGTCTGTCGCGGGAGAGCCCGAGGCTCTGTGGCGGCCCGCGAGGAGGGA 1391
Db 1381 CACAGACCCCGTCCCTGTGTGAGTGTCTCCGCGCAGCAGACGCCCTGCGAGGTGTA 1440
QY 1392 CACAGACCCCGTCCCTGTGTGAGTGTCTCCGCGCAGCAGACGCCCTGCGAGGTGTA 1451
Db 1441 CGGCTTCTGTCGCGGCCCTGCTCGCGGCCGTGTGTCGCCAGCCCTCTGGGCTCCAGGCA 1500
QY 1452 CGGCTTCTGTCGCGGCCCTGCTCGCGGCCGTGTGTCGCCAGCCCTCTGGGCTCCAGGCA 1511
Db 1501 CAAGACCGCCGCTCTCTCAGACACCAAGAGTTCATCTCTCTGGGAGAGCATGCCAA 1560
QY 1512 CAAGACCGCCGCTCTCTCAGACACCAAGAGTTCATCTCTCTGGGAGAGCATGCCAA 1571
Db 1561 GCTCTCGCTCAGAGCTGAGTGTGAAGATGAGCGTGGGAGTCCGCTTGGCTGCCGAG 1620
QY 1572 GCTCTCGCTCAGAGCTGAGTGTGAAGATGAGCGTGGGAGTCCGCTTGGCTGCCGAG 1631
Db 1621 GAGCCCGAGGGTGGCTGTGTTCGCGCCGCGAGACACCGCTTGGGTGAGGAGATCCTGGC 1680
QY 1632 GAGCCCGAGGGTGGCTGTGTTCGCGCCGCGAGACACCGCTTGGGTGAGGAGATCCTGGC 1691
Db 1681 CAAGTCTCTGACCTGGCTGATGATGTGTGAGTGTGAGTGTGAGTGTCTCTTTT 1740
QY 1692 CAAGTCTCTGACCTGGCTGATGATGTGTGAGTGTGAGTGTGAGTGTCTCTTTT 1751
Db 1741 TGTCACGAGACACAGTTTCAAAAGACAGCTCTTTTCTACCGGAGAGATGTCTGGAG 1800
QY 1752 TGTCACGAGACACAGTTTCAAAAGACAGCTCTTTTCTACCGGAGAGATGTCTGGAG 1811
Db 1801 CAAGTGTCAAGAGCATTTGAATCAGACAGCACTTGAAGAGGGTGCAGTTCGGGAGCTGTC 1860
QY 1812 CAAGTGTCAAGAGCATTTGAATCAGACAGCACTTGAAGAGGGTGCAGTTCGGGAGCTGTC 1871
Db 1861 GGAAGCAGAGTCTAGGAGCATTCGGGAGAGCCAGCCCGCTCTGCTGACG 1909
QY 1872 GGAAGCAGAGTCTAGGAGCATTCGGGAGAGCCAGCCCGCTCTGCTGACG 1920

RESULT 8

US-08-854-050-224 STANDARD; DNA; UNC; 4015 BP.
xxxxxx
Sequence 224, Application US/08854050
Sequence 224, Application US/08854050
GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Novel telomerase
CC NUMBER OF SEQUENCES: 225
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002930US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 224:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4015 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cdna
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 56..3454 /product= "hprt"
CC OTHER INFORMATION: /note= "human telomerase reverse
CC OTHER INFORMATION: transcriptase (hprt) catalytic protein
CC OTHER INFORMATION: component"
SQ SEQUENCE 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 OTHER.
Query Match 99.4%; Score 1909; DB 21; Length 4015;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GCAGCGCTGCGTCTGCTGCGCACGTGGAGAGCCCTGCGCCCGGCGCCACCCCGCGGATGCC 60
Qy 12 GCAGCGCTGCGTCTGCTGCGCACGTGGAGAGCCCTGCGCCCGGCGCCACCCCGCGATGCC 71
Db 61 GCGCGCTCCCGCTCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Qy 72 GCGCGCTCCCGCTCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 131
Db 121 GCGCGTGGCGACGTTGCTGCGCGCGCTGCGGCGCCACAGGCGCTGCGGCTGCTGCGAGCGG 180
Qy 132 GCGCGTGGCGACGTTGCTGCGCGCGCTGCGGCGCCACAGGCGCTGCGGCTGCTGCGAGCGG 191
Db 181 GGAACCGGCGGCTTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Qy 192 GGAACCGGCGGCTTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 251
Db 241 AGCGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Qy 252 AGCGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 311
Db 301 CCGAGTGTGCAAGAGCTGTGCGAGCGCGCGCGCGAGAGAGCTGCTGCGCTTCCGCTTCCG 360
Qy 312 CCGAGTGTGCAAGAGCTGTGCGAGCGCGCGCGCGAGAGAGCTGCTGCGCTTCCGCTTCCG 371
Db 361 GCTGCTGAGAGCGGCG 420
Qy 372 GCTGCTGAGAGCGGCG 431
Db 421 CCG 480
Qy 432 CCG 491
Db 481 CCGCGCTGGCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 492 CCGCGTGGCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551
Db 541 GCGTCCAGCTCGGCTACCAAGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Qy 552 GCGTCCAGCTCGGCTACCAAGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 611
Db 601 TCAGCG 660
Qy 612 TCAGCG 671
Db 661 CTGGAACCATAGCTGAGGAGCGCGGCTGCGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCG 720
Qy 672 CTGGAACCATAGCTGAGGAGCGCGGCTGCGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCG 731
Db 721 GAGGCGCGGCGAGTGCAGCGCGAGTGTGCGGTTGCCAGAGCGCGCGCGCGCGCGCGCGCG 780
Qy 732 GAGGCGCGGCGAGTGCAGCGCGAGTGTGCGGTTGCCAGAGCGCGCGCGCGCGCGCGCGCG 791
Db 781 TCGCGCTGAGCGGAGCGGAGCG 840
Qy 792 TCGCGCTGAGCGGAGCGGAGCG 851
Db 841 GCGTGGACCGAGTACCGTGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy 852 GCGTGGACCGAGTACCGTGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911
Db 901 CACCTCTTTGGAGGTTGCGCTCTGCGCACGCGCGCACTCCCAACCACTGCGTGGCGCGCGCA 960
Qy 912 CACCTCTTTGGAGGTTGCGCTCTGCGCACGCGCGCACTCCCAACCACTGCGTGGCGCGCGCA 971
Db 961 GCACACG 1020
Qy 972 GCACACG 1031
Db 1021 CCGGCTGTACGCGAGCAAGCACTTCTACTCTCTACTCTCTACTCTCTACTCTCTACTCTCTCT 1080
Qy 1032 CCGGCTGTACGCGAGCAAGCACTTCTACTCTCTACTCTCTACTCTCTACTCTCTACTCTCTCT 1091

Db 1081 GCGCTCTTCTCTACTCTAGCTCTCTGAGGCGCCAGCGCTGACTGGCGCTCGGAGGCTCGTGGA 1140
Qy 1092 GCGCTCTTCTCTACTCTAGCTCTCTGAGGCGCCAGCGCTGACTGGCGCTCGGAGGCTCGTGGA 1151
Db 1141 GACCATCTTCTGGGTTCCAGGCGCTGGATGCCAGGAGCTCCCGCGAGGTTGCCCGCGCT 1200
Qy 1152 GACCATCTTCTGGGTTCCAGGCGCTGGATGCCAGGAGCTCCCGCGAGGTTGCCCGCGCT 1211
Db 1201 GCGCCAGCGCTACTGCAAAATGCGGCCCTGTCTTCTGGAGCTGCTTGGGAACCAACGCGCA 1260
Qy 1212 GCGCCAGCGCTACTGCAAAATGCGGCCCTGTCTTCTGGAGCTGCTTGGGAACCAACGCGCA 1271
Db 1261 GTGCGCTACGGGTGCTCTCAAGACGCACTGCGCGCTGGAGCTCGGCTCAACCGCGAGC 1320
Qy 1272 GTGCGCTACGGGTGCTCTCAAGACGCACTGCGCGCTGGAGCTCGGCTCAACCGCGAGC 1331
Db 1321 AGCGGCTGTCTGCCGCGGAGAACCCAGGCGCTGTGTGGCGCGCCCGAGGAGGAGGA 1380
Qy 1332 AGCGGCTGTCTGCCGCGGAGAACCCAGGCGCTGTGTGGCGCGCCCGAGGAGGAGGA 1391
Db 1381 CACAGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Qy 1392 CACAGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1451
Db 1441 CGGCTTGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Qy 1452 CGGCTTGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511
Db 1501 CAAGAAACCGCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAAGCATGCCAA 1560
Qy 1512 CAAGAAACCGCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAAGCATGCCAA 1571
Db 1561 GCTCTGCTGCGAGAGCTGAGCTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTTGGCTGCGAG 1620
Qy 1572 GCTCTGCTGCGAGAGCTGAGCTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGCGAG 1631
Db 1621 GAGCCAGGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Qy 1632 GAGCCAGGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1691
Db 1681 CAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Qy 1692 CAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1751
Db 1741 TGTCAAGGAGACCGTGTTCAAAAGAACAGCGCTTTTTTACCGGAAGAGTGTCTGGAG 1800
Qy 1752 TGTCAAGGAGACCGTGTTCAAAAGAACAGCGCTTTTTTACCGGAAGAGTGTCTGGAG 1811
Db 1801 CAAGTTCGAAGCATTGAATCAGACAGCACTTGAAGAGGCTGAGCTGCGGAGCTGTC 1860
Qy 1812 CAAGTTCGAAGCATTGAATCAGACAGCACTTGAAGAGGCTGAGCTGCGGAGCTGTC 1871
Db 1861 GGAAGCAGAGTCAAGGAGCATCGGGAAGCCAGCGCCCGCTGCTGAGC 1909
Qy 1872 GGAAGCAGAGTCAAGGAGCATCGGGAAGCCAGCGCCCGCTGCTGAGC 1920

RESULT 9
ID US-09-026-981-35 STANDARD; DNA; UNC; 4023 BP.
AC xxxxxx
DT
DE Sequence 35, Application US/09026981
CC Sequence 35, Application US/09026981
CC GENERAL INFORMATION:
CC APPLICANT: Counter, Christopher M.
CC APPLICANT: Meyerson, Matthew
CC APPLICANT: Weinberg, Robert A.
CC TITLE OF INVENTION: Telomerase Catalytic Subunit Gene and
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Militia Drive
CC CITY: Lexington

CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US 60/026,981
CC FILING DATE: 20-FEB-1998
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/064,322
CC FILING DATE: 30-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/055,762
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/054,549
CC FILING DATE: 01-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/047,151
CC FILING DATE: 20-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/038,750
CC FILING DATE: 20-FEB-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WH197-11p4AM
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 781-861-6240
CC TELEFAX: 781-861-9540
CC INFORMATION FOR SEQ ID NO: 35:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4023 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 4023 BP: 668 A; 1363 C; 1277 G; 715 T; 0 OTHER.

Query Match 99.3%; Score 1906; DB 23; Length 4023;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1908; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 3 GGCAGCGCTGGGTCTCTGCTGCGCAGCTGGGAAGCCCTGGCCCGGCCACCCCGCGATGC 62
|||||
Qy 11 GGCAGCGCTGGGTCTCTGCTGCGCAGCTGGGAAGCCCTGGCCCGGCCACCCCGCGATGC 70
|||||
Db 63 CGCGCGCTCCCGCTGCGAGCGCGTGCCTCCCTGCTGCGCAGCCACTACCGCGAGTGC 122
|||||
Qy 71 CGCGCGCTCCCGCTGCGAGCGCGTGCCTCCCTGCTGCGCAGCCACTACCGCGAGTGC 130
|||||
Db 123 TGCCGCTGGCCACGTTCTGCGGCGCTGGGGCCCCAGGGCTGGCGGTGCTGAGCGCG 182
|||||
Qy 131 TGCGCGTGGCCACGTTCTGCGGCGCTGGGGCCCCAGGGCTGGCGGTGCTGAGCGCG 190
|||||
Db 183 GGGACCCGGCGCTTTCGCGCGCTGGTGGCCAGTGCCTGGTGGCTGCCCTGGGACG 242
|||||
Qy 191 GGGACCCGGCGCTTTCGCGCGCTGGTGGCCAGTGCCTGGTGGCTGCCCTGGGACG 250
|||||
Db 243 CACGCGCCCGCCCGCCCGCCCTCTTCGCGCAGGTCTCTGCGTGAAGAGCTGGTGG 302
|||||
Qy 251 CACGCGCCCGCCCGCCCGCCCTCTTCGCGCAGGTCTCTGCGTGAAGAGCTGGTGG 310
|||||
Db 303 CCGAGTCTGAGAGGTGTGCGAGCGCGCGGCGAAGAGTGTGCTGGCTTGGCTTCG 362
|||||
Qy 311 CCGAGTCTGAGAGGTGTGCGAGCGCGCGGCGAAGAGTGTGCTGGCTTGGCTTCG 370
|||||
Db 363 CGCTGCTGAGCGGCGCGCGGCGCGCCCGCGAGGCTTACCACAGCGCTGGCGAGCT 422
|||||
Qy 371 CGCTGCTGAGCGGCGCGCGGCGCGCCCGCGAGGCTTACCACAGCGCTGGCGAGCT 430
|||||

Db 423 ACCTGCCCAACACAGGTGACCGACCGCACTGCGGGGGAGCGGGGCGTGGGGGCTGCTGTGC 482
|||||
Qy 431 ACCTGCCCAACACAGGTGACCGACCGCACTGCGGGGGAGCGGGGCGTGGGGGCTGCTGTGC 490
|||||
Db 483 GCCCGGTGGCGAGCAGTGTGTTTACCTGCTGCGACGCTGCGCGCTCTTTTGTGTGG 542
|||||
Qy 491 GCCCGGTGGCGAGCAGTGTGTTTACCTGCTGCGACGCTGCGCGCTCTTTTGTGTGG 550
|||||
Db 543 TGGTCCCAAGTGGCTTACCAAGTGTGCGGGCGCGCGCTGTACCAAGTGGCGCTGCCA 602
|||||
Qy 551 TGGTCCCAAGTGGCTTACCAAGTGTGCGGGCGCGCGCTGTACCAAGTGGCGCTGCCA 610
|||||
Db 603 CTAGAGCCCGCGCCCGCCACACGCTAGTGAAGCCCGGAGCGTGTGGATGCGAAGCGG 662
|||||
Qy 611 CTAGAGCCCGCGCCCGCCACACGCTAGTGAAGCCCGGAGCGTGTGGATGCGAAGCGG 670
|||||
Db 663 CCTGGAACCATAGCTCAGGAGCGCGGGTCCCGCTGGCGCTGCCAGCCCGGGTGGCA 722
|||||
Qy 671 CCTGGAACCATAGCTCAGGAGCGCGGGTCCCGCTGGCGCTGCCAGCCCGGGTGGCA 730
|||||
Db 723 GGAGCGCGGGGCGAGTGCACAGCCGAAAGTGTGCGTTGCCCAAGAGCGCCAGGCGTGGCG 782
|||||
Qy 731 GGAGCGCGGGGCGAGTGCACAGCCGAAAGTGTGCGTTGCCCAAGAGCGCCAGGCGTGGCG 790
|||||
Db 783 CTGCCCCGTAGCGCGGAGCGGCCCGTGGGCGAGGGGTCTTGGGCGCCACCGCGGAGGA 842
|||||
Qy 791 CTGCCCCGTAGCGCGGAGCGGCCCGTGGGCGAGGGGTCTTGGGCGCCACCGCGGAGGA 850
|||||
Db 843 CGCGTGGACCGAGTGACCGTGGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 902
|||||
Qy 851 CGCGTGGACCGAGTGACCGTGGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 910
|||||
Db 903 CCACCTCTTTGGAGGTGCGCTCTGTGCGACGCGCCCACTCCACCCCATCCGTGGGCGGCC 962
|||||
Qy 911 CCACCTCTTTGGAGGTGCGCTCTGTGCGACGCGCCCACTCCACCCCATCCGTGGGCGGCC 970
|||||
Db 963 AGCACACGCGGGCGCCCGCCATCCACATCGGCGGCCACACGTCCTTGGGACACGCTTGTG 1022
|||||
Qy 971 AGCACACGCGGGCGCCCGCCATCCACATCGGCGGCCACACGTCCTTGGGACACGCTTGTG 1030
|||||
Db 1023 CCGCGGTGTAGCGCGAGCAACGACTTCTCTACTCTCCTAGGCGGACAGGAGCAGTGC 1082
|||||
Qy 1031 CCGCGGTGTAGCGCGAGCAACGACTTCTCTACTCTCCTAGGCGGACAGGAGCAGTGC 1090
|||||
Db 1083 GGCCCTCTTCTACTACTCAGCTCTCTGAGGCGCCAGCTGACTGGGCGCTCGAGGCTCGTGG 1142
|||||
Qy 1091 GGCCCTCTTCTACTACTCAGCTCTCTGAGGCGCCAGCTGACTGGGCGCTCGAGGCTCGTGG 1150
|||||
Db 1143 AGACCATCTTTCTGGGTTCAGGCGCCCTGGATGCGAGGAGTCCCGCGAGTGTGCCCGGCC 1202
|||||
Qy 1151 AGACCATCTTTCTGGGTTCAGGCGCCCTGGATGCGAGGAGTCCCGCGAGTGTGCCCGGCC 1210
|||||
Db 1203 TGCCCGCAGCGCTACTGCGCAATGCGGCCCTGTCTTCTGGAGCTGCTTGGGNAACACGCGC 1262
|||||
Qy 1211 TGCCCGCAGCGCTACTGCGCAATGCGGCCCTGTCTTCTGGAGCTGCTTGGGAACACGCGC 1270
|||||
Db 1263 AGTGCCTTACGCGGTGCTCTCTCAAGACGACTGCCCGCTCGAGCTGCGGTGCTACCCCGAG 1322
|||||
Qy 1271 AGTGCCTTACGCGGTGCTCTCTCAAGACGACTGCCCGCTCGAGCTGCGGTGCTACCCCGAG 1330
|||||
Db 1323 CAGCGGTGTGTGTGCCCGGAGAACCCAGGCTCTGTGGCGGCCCGCCCGAGGAGGAGG 1382
|||||
Qy 1331 CAGCGGTGTGTGTGCCCGGAGAACCCAGGCTCTGTGGCGGCCCGCCCGAGGAGGAGG 1390
|||||
Db 1383 ACACAGACCCCGCTGCGCTGTGTGAGTGTCTCCGCGACGACAGACGCGCTTGGAGTGT 1442
|||||
Qy 1391 ACACAGACCCCGCTGCGCTGTGTGAGTGTCTCCGCGACGACAGCAGCGCTTGGAGTGT 1450
|||||
Db 1443 ACAGCTTGTGCGGCGCTGCTGCGCGGTGTGTGCGCGCGCTTGGGCTTCCAGGC 1502
|||||
Qy 1451 ACAGCTTGTGCGGCGCTGCTGCGCGGTGTGTGCGCGCGCTTGGGCTTCCAGGC 1510
|||||

Db 659 CTGGAACCATAGCTCAGGAGGCGGGTCCCTCGGCTGCGC-TGCCAGAGCCCGGGTGCAG 717
Qy 672 CTGGAACCATAGCTCAGGAGGCGGGTCCCTCGGCTGCGCAGCCCGGGTGCAG 731
Db 718 GAGCGCGGGGAGTGCAGCCAGAGTCTGCCGTTGCCAAGAGGCGCCAGCGTGGCG 777
Qy 732 GAGCGCGGGGAGTGCAGCCAGAGTCTGCCGTTGCCAAGAGGCGCCAGCGTGGCG 791
Db 778 TGCCCTCAGCGGAGCGAGCCCGTTGGCAGGGGCTCTGGGCGCCACCGGGCAGGAC 837
Qy 792 TGCCCTCAGCGGAGCGAGCCCGTTGGCAGGGGCTCTGGGCGCCACCGGGCAGGAC 851
Db 838 GCCTGGACCGAGTGACCGTGGTCTGTGGTGTGACCTGCGACAGCCCGCGAAGAGC 897
Qy 852 GCGTGGACCGAGTGACCGTGGTCTGTGGTGTGACCTGCGACAGCCCGCGAAGAGC 911
Db 898 CACCTCTTTGGAGGGTGGCTCTCTGGCAGCGCCACCTCCACCCATCCGTGGCGCGCA 957
Qy 912 CACCTCTTTGGAGGGTGGCTCTCTGGCAGCGCCACCTCCACCCATCCGTGGCGCGCA 971
Db 958 GCACACGCGGCGCCCATCACATCGCGCCACACGCTCC-TGGGACAGCCTTGTC 1016
Qy 972 GCACACGCGGCGCCCATCACATCGCGCCACACGCTCCCTGGGACAGCCTTGTC 1031
Db 1017 CCGGTGTACGCGGAGACCAAGCACTTCTCTACTCTCAGCGGACAAGNA-CACTGCGN 1075
Qy 1032 CCGGTGTACGCGGAGACCAAGCACTTCTCTACTCTCAGCGGACAAGNAGCACTGCG 1091
Db 1076 CCGTCTTCTACTACTATATCTGAGCGCCAGCTGACTGCGTTCGGGAGGTCTG 1135
Qy 1092 GCGTCTTCTACTACTACTCTCTGAGCGCCAGCGTACTGCGCTCGGAGTCTG 1151
Db 1136 AGACANTCTTTCTGTTCCAGGCTTGATGCGCAGG-ATTCCCGCAGGTTGCCCGGCT 1194
Qy 1152 GACCATCTTTCTGGTTCCAGGCTTGATGCGCAGGACTCCCGCAGGTTGCCCGGCT 1211
Db 1195 GCCCGAGCGTACTGGCAAAATCGCGCCCTGTTCTGAGTGTCTGTTGGAAACCGCGCA 1254
Qy 1212 GCCCGAGCGTACTGGCAAAATCGCGCCCTGTTCTGAGTGTCTGTTGGAAACCGCGCA 1271
Db 1255 GTGCGGCTACGGGTTGCTCTCAGACGCACTGCCGCTGCGAGTGGGTACCCCGAGC 1314
Qy 1272 GTGCGGCTACGGGTTGCTCTCAGACGCACTGCCGCTGCGAGTGGGTACCCCGAGC 1331
Db 1315 AGCGGTGTCTGTGCGCGGAGAACCCAGGCTCTGTGCGCGCCCGCGGAGGAGGA 1374
Qy 1332 AGCGGTGTCTGTGCGCGGAGAACCCAGGCTCTGTGCGCGCCCGCGGAGGAGGA 1391
Db 1375 ACACAGACCCCGTGTGCTGTGAGTGTCTCCCGCAGCAGCAGCCCTGGCAGGTGT 1434
Qy 1392 -CACAGACCCCGTGTGCTGTGAGTGTCTCCCGCAGCAGCAGCCCTGGCAGGTGT 1450
Db 1435 ACGGGCTGTGCGGCGCTGCTGCGCGCGGTGGTCCCGCAGGCTCTGGGGCTCCAGGC 1494
Qy 1451 ACGGGCTGTGCGGCGCTGCTGCGCGCGGTGGTCCCGCAGGCTCTGGGGCTCCAGGC 1510
Db 1495 ACAAGAACCGCGCTCTCTAGGAACACACAGAGTTCATCTCCCTGGGAGCATGCCA 1554
Qy 1511 ACAAGAACCGCGCTCTCTAGGAACACACAGAGTTCATCTCCCTGGGAGCATGCCA 1570
Db 1555 AGCTCTCGCTGCAGGAGTGCAGTGGAAAGATGAGGCTCGGGAGTGCCTTGGCTGCSCA 1614
Qy 1571 AGCTCTCGCTGCAGGAGTGCAGTGGAAAGATGAGGCTCGGGAGTGCCTTGGCTGCSCA 1630
Db 1615 GGAGCCGAGGGTGGCTGTGTTCCGCGCGCAGACAGCAGCTGCGTAGAGATCCTGG 1674
Qy 1631 GGAGCCGAGGGTGGCTGTGTTCCGCGCGCAGACAGCAGCTGCGTAGAGATCCTGG 1690
Db 1675 CCAAGTCTCTGACGTGGTGTAGTGTGTAGTGTGTAGTGTGTAGTGTGTCTCTTTT 1734
Qy 1691 CCAAGTCTCTGACGTGGTGTAGTGTGTAGTGTGTAGTGTGTAGTGTGTCTCTTTT 1750

Db 1735 ATGTCAGGAGACACAGTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGA 1794
Qy 1751 ATGTCAGGAGACACAGTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGA 1810
Db 1795 GCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGTCAGCTGGGGAGCTGT 1854
Qy 1811 GCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGTCAGCTGGGGAGCTGT 1870
Db 1855 CGGAAGCAGAGGTCAGGCAGCATCGGAAGCCAGCGCCCGCTCTGCTGACG 1904
Qy 1871 CGGAAGCAGAGGTCAGGCAGCATCGGAAGCCAGCGCCCGCTCTGCTGACG 1920
RESULT 11
ID US-08-851-843-173 STANDARD: DNA; UNC; 4029 BP.
AC xxxxxx
DT
DE Sequence 173, Application US/08851843
CC Sequence 173, Application US/08851843
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Novel Telomerase
CC NUMBER OF SEQUENCES: 223
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA: US/08/851,843
CC APPLICATION NUMBER: US/08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002930US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 173:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4029 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY:

CC LOCATION: 1..4029 /note= "preliminary sequence for
CC OTHER INFORMATION: human TRT cDNA insert of
CC OTHER INFORMATION: plasmid pGRN121"
SQ SEQUENCE 4029 BP; 687 A; 1342 C; 1255 G; 720 T; 25 OTHER.

Query Match 83.1%; Score 1596; DB 21; Length 4029;
Best Local Similarity 95.7%; Pred. No. 0.00e+00;
Matches 1828; Conservative 0; Mismatches 75; Indels 7; Gaps 6;

Db 1 GCAGCGCTGCGTCTGCGCACGTGGGAAGCCCTGGCCCGGCGCCACCCCGCGCATGCC 60
Qy 12 GCAGCGCTGCGTCTGCGCACGTGGGAAGCCCTGGCCCGGCGCCACCCCGCGCATGCC 71
Db 61 GCGCGCTCCCGCTCCGAGCGCTGCGTCTGCTGGCGAGCCACTACCGGAGGTGCT 120
Qy 72 GCGCGCTCCCGCTCCGAGCGCTGCGTCTGCTGGCGAGCCACTACCGGAGGTGCT 131
Db 121 GCGCGTGGCGACGTTGCTGGCGGCTTGGGCGCCCGAGGCTGGCGGTGGTGCAGCGGG 180
Qy 132 GCGCGTGGCGACGTTGCTGGCGGCTTGGGCGCCCGAGGCTGGCGGTGGTGCAGCGGG 191
Db 181 GGACCGCGCGGCTTTCGCGCGGCTGGTGGCCGANTGNTGGTGGTGGTGGTGGTGGTGG 240
Qy 192 GGACCGCGCGGCTTTCGCGCGGCTGGTGGCCGAGTGGCTGGTGGTGGTGGTGGTGG 251
Db 241 ANGGCGCCCGCGCGCGCGCTTCTCCGCGAGGTGCTGCTGCTGAANGANTGCTGGCG 300
Qy 252 ACGGCGCGCGCGCGCGCGCTTCTCCGCGAGGTGCTGCTGCTGAANGANTGCTGGCG 311
Db 301 CCGAGTGTGCGANAGCTGTGGGAGCGGCGCGGAGGAGTGGTGGTGGTGGTGGTGGTGG 360
Qy 312 CCGAGTGTGCGAGGCTGTGGAGCGGCGGCGGAGGAGTGGTGGTGGTGGTGGTGGTGG 371
Db 361 GCTGTGTGACGGGGCGCGGGGGCGCCCGAGGCGCTTACACACAGCGTGGCGAGCTA 420
Qy 372 GCTGTGTGACGGGGCGCGGGGGCGCCCGAGGCGCTTACACACAGCGTGGCGAGCTA 431
Db 421 CTGCGCCACACAGCGTACCGACGCTGCGGGGAGCGGGGCTGCTGCTGGCG 480
Qy 432 CTGCGCCACACAGCGTACCGACGCTGCGGGGAGCGGGGCTGCTGCTGGCG 491
Db 481 CCGCGTGGCGAGAGCTGCTGGTTCACCTGCTGGCAGCGTGGCGGNTTGTGCTGGT 540
Qy 492 CCGCGTGGCGAGAGCTGCTGGTTCACCTGCTGGCAGCGTGGCGGCTTGTGCTGGT 551
Db 541 GNTTCCAGCTCGCGCTACCAAGTGTGGGGCGCGCTGTACCAAGCTGGCGTGCNAC 600
Qy 552 GCTCCAGCTCGCGCTACCAAGTGTGGGGCGCGCTGTACCAAGCTGGCGTGCNAC 611
Db 601 TCAGGCGGGCGCGCGCACAGCTTANTGGACCGAANGCTGCTGGGAT--CCAAAGGCG 658
Qy 612 TCAGGCGGGCGCGCGCACAGCTAGTGGACCGAAGGCTCTGGGATGCGAAGGCG 671
Db 659 CTGGAACCATAGCTGAGGAGCGGGGTCCCTCTGGGC--TGCCAGCCCGGGTGCAG 717
Qy 672 CTGGAACCATAGCTGAGGAGCGGGGTCCCTCTGGGCCTGCCAGCCCGGGTGCAG 731
Db 718 GAGGCGGGGGCGAGTGCAGCGCAAGTGTGCGGCTGCGGAGAGCGCGGCTGGCG 777
Qy 732 GAGGCGGGGGCGAGTGCAGCGCAAGTGTGCGGCTGCGGAGAGCGCGGCTGGCG 791
Db 778 TGCCCTCTGAGCGCGGAGCGCGCGTGGGAGGGGCTTGGGCGCCACCGGGGAGGAC 837
Qy 792 TGCCCTCTGAGCGCGGAGCGCGCGTGGGAGGGGCTTGGGCGCCACCGGGGAGGAC 851
Db 838 GCTTGACCGAGTACCGTGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 897
Qy 852 GCGTGGACCGAGTACCGTGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 911
Db 898 CACCTCTTTGAGGTGGCTCTCTGGCAGCGCGCCACTCCACCCATCGTGGGCGCGCA 957
Qy 912 CACCTCTTTGAGGTGGCTCTCTGGCAGCGCGCCACTCCACCCATCGTGGGCGCGCA 971

RESULT 12

ID US-08-912-951-6 STANDARD; DNA; UNC; 4200 BP.

AC xxxxxx

DT

DE Sequence 6, Application US/08912951

CC Sequence 6, Application US/08912951

GENERAL INFORMATION:

CC

CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC THERAPEUTIC METHODS
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-0026000S
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4200 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 4200 BP; 525 A; 1187 C; 1138 G; 644 T; 706 OTHER.

Query Match 78.0%; Score 1497; DB 22; Length 4200;
Best Local Similarity 94.0%; Pred. No. 0.00e+00;
Matches 1630; Conservative 0; Mismatches 0; Indels 104; Gaps 1;

Db 2257 GGCAGCGTGGCTCTGCTGCGCAGTGGGAAGCCCTGGCCCGGCCACCCCGCGATGC 2316
|||||
Qy 11 GGCAGCGTGGCTCTGCTGCGCAGTGGGAAGCCCTGGCCCGGCCACCCCGCGATGC 70
|||||

Db 2317 GCGCGCTCCCGCTGCGGAGCCGTGGCTCCCTGCTGCGCAGCAGCTACCGCGAGGTGC 2376
|||||
Qy 71 GCGCGCTCCCGCTGCGGAGCCGTGGCTCCCTGCTGCGCAGCAGCTACCGCGAGGTGC 130
|||||

Db 2377 TGCCGCTGGCCACAGTTGCTGCGGGGGCCCTGGGGCCCCAGAGGCTGGCGGCTGGTGCAGGGCG 2436
|||||
Qy 131 TGCCGCTGGCCACAGTTGCTGCGGGGGCCCTGGGGCCCCAGAGGCTGGCGGCTGGTGCAGGGCG 190
|||||
Db 2437 GGCACCGGGCGGCTTCCCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGTGGTGGTGGTGGTGGTGG 2496
|||||
Qy 191 GGCACCGGGCGGCTTCCCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGTGGTGGTGGTGGTGGTGG 250
|||||
Db 2497 CAGGGCGCCCGCGCGCGCGCCCTTCCTTCCGCCAGGTGGGCCCTCCCGGGGCTGGCGTCC 2556
|||||
Qy 251 CAGGGCGCCCGCGCGCGCGCCCTTCCTTCCGCCAGGTG----- 288
|||||
Db 2557 GGCTGGGGTTGAGGGGGCGGGGGGAACACGCGACATCGGAGAGCAGCGCAGGCGACT 2616
|||||
Qy 288 ----- 288
|||||
Db 2617 CAGGGCGCTTCCCGCGCAGGTGCTGCTGGAAGAGCTGGTGGCCCGAGTGTGCAGAG 2676
|||||
Qy 289 -----TCCCTGCCCTGAAGAGCTGGTGGCCCGAGTGTGCAGAG 326
|||||
Db 2677 GCTGTGCGAGCGGGCGGCGAAGACGTGCTGGCTTTCGGCTTCGCGCTTCGCGTGTGACGGGCG 2736
|||||
Qy 327 GCTGTGCGAGCGGGCGGCGAAGACGTGCTGGCTTTCGGCTTCGCGTGTGCGTGTGACGGGCG 386
|||||
Db 2737 CCGGGGGGCCCCCGGAGGCTTTCACACAGCGTGGCAGTACTGCCCCAACACGGT 2796
|||||
Qy 387 CCGGGGGGCCCCCGGAGGCTTTCACACAGCGTGGCAGTACTGCCCCAACACGGT 446
|||||
Db 2797 GACCGAGCACTGGGGGGGAGCGGGGCTGGGGGCTGCTGCGCGCGCTGGCGGCGTGGCGCAGCA 2856
|||||
Qy 447 GACCGAGCACTGGGGGGGAGCGGGGCTGGGGGCTGCTGCGCGCGTGGGGCGCAGCA 506
|||||
Db 2857 CGTGTGTTCACTGCTGGCACGCTGCGCGCTCTTTGTGCTGTGTGCTGCCAGTCCGCTGCGC 2916
|||||
Qy 507 CGTGTGTTCACTGCTGGCACGCTGCGCGCTCTTTGTGCTGTGTGCTGCCAGTCCGCTGCGC 566
|||||
Db 2917 CTACAGGTGTGGGGCGCGCTGTACACAGCTGGCGCTGCCACTCAGCGCCCGGCGCCCC 2976
|||||
Qy 567 CTACAGGTGTGGGGCGCGCTGTACACAGCTGGCGCTGCCACTCAGCGCCCGGCGCCCC 626
|||||
Db 2977 GGCACAGCTAGTGGAGCCCGAAGCGCTGGGATGGGAAGCGGCTGGAAACCATAGCGT 3036
|||||
Qy 627 GCCACAGCTAGTGGAGCCCGAAGCGCTGGGATGGGAAGCGGCTGGAAACCATAGCGT 686
|||||
Db 3037 CAGGGAGCGGGTCCCTGGCGCTGCCAGCCCGGGTGGCAGGAGGCGCGGGGCGAG 3096
|||||
Qy 687 CAGGGAGCGGGTCCCTGGCGCTGCCAGCCCGGGTGGCAGGAGGCGCGGGGCGAG 746
|||||
Db 3097 TGCCAGCCGAAGTCTGCGCTTGCCCAAGAGCCCGAGCGTGGCGCTGCCCTGAGCCGGA 3156
|||||
Qy 747 TGCCAGCCGAAGTCTGCGCTTGCCCAAGAGCCCGAGCGTGGCGCTGCCCTGAGCCGGA 806
|||||
Db 3157 GCGAGCCCGTGGGAGGGGCTCTGGGCCACCGCGGCGAGCAGCGTGGACCGAGTGA 3216
|||||
Qy 807 GCGAGCCCGTGGGAGGGGCTCTGGGCCACCGCGGCGAGCAGCGTGGACCGAGTGA 866
|||||
Db 3217 CCGTGGTTCTGTGTGTGCTCACTGCGACACCGCGGAGAGCCACCTCTTTGGAGGG 3276
|||||
Qy 867 CCGTGGTTCTGTGTGTGCTCACTGCGACACCGCGGAGAGCCACCTCTTTGGAGGG 926
|||||
Db 3277 TGCGCTCTGTGGACCGCGCCACTCCACCCATCGTGGGCGGCGCAGCACCACCGCGGCGCC 3336
|||||
Qy 927 TGCGCTCTGTGGACCGCGCCACTCCACCCATCGTGGGCGGCGCAGCACCACCGCGGCGCC 986
|||||
Db 3337 CCATCCACATCGGGGCCACCGTCCCTGGGACACGCTTGTCCCGCGGTGTACGCCGA 3396
|||||
Qy 987 CCATCCACATCGGGGCCACCGTCCCTGGGACACGCTTGTCCCGCGGTGTACGCCGA 1046
|||||
Db 3397 GACCAACACTTCTCTACTCTCCTCAGCGCAGAGGAGCAGCTGGGGCCCTCTCTCTACT 3456
|||||
Qy 1047 GACCAACACTTCTCTACTCTCCTCAGCGCAGAGGAGCAGCTGGGGCCCTCTCTCTACT 1106
|||||

Db 3457 CAGCTCTGAGGCCAGCCTGACTGGCGCTCGGAGGCTCGTGAGAGCACCATTCTTCTGGG 3516
QY 1107 CAGCTCTGAGGCCAGCCTGACTGGCGCTCGGAGGCTCGTGAGAGCACCATTCTTCTGGG 1166
Db 3517 TTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCTGCGCCAGCGCTACTG 3576
QY 1167 TTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCTGCGCCAGCGCTACTG 1226
Db 3577 GCAATATGGGCCCTGTTTCTTGAGCTGCTTGGAGTACACGCGAGTCCCTTACGGGGT 3636
QY 1227 GCAATATGGGCCCTGTTTCTTGAGCTGCTTGGAGTACACGCGAGTCCCTTACGGGGT 1286
Db 3637 GCTCTCAAGACGCTGCTGCGCTCGGAGTCTCGGTCTACCCAGCAGCGGTGTGTGTC 3696
QY 1287 GCTCTCAAGACGCTGCTGCGCTCGGAGTCTCGGTCTACCCAGCAGCGGTGTGTGTC 1346
Db 3697 CCGGAGAGGCCCGGCTGCTGGCGGCCCGCCAGGAGGAGACAGACCCCGCTCG 3756
QY 1347 CCGGAGAGGCCCGGCTGCTGGCGGCCCGCCAGGAGGAGACAGACCCCGCTCG 1406
Db 3757 CTTGTGAGCTGCTCCGCCAGCAGCAGCCCTGCGAGGTGTACGGCTTCTGTCGGGC 3816
QY 1407 CTTGTGAGCTGCTCCGCCAGCAGCAGCCCTGCGAGGTGTACGGCTTCTGTCGGGC 1466
Db 3817 CTGCTGCGCGGCTGCTGCGCCAGGCTCTGGGGCTCCAGGACACAGACGCCGCTT 3876
QY 1467 CTGCTGCGCGGCTGCTGCGCCAGGCTCTGGGGCTCCAGGACACAGACGCCGCTT 1526
Db 3877 CTTAGGAGACCAAGAGTTCATCTCCTGGGAGGATGCAAGTCTCGCTCAGGA 3936
QY 1527 CTTAGGAGACCAAGAGTTCATCTCCTGGGAGGATGCAAGTCTCGCTCAGGA 1586
Db 3937 GCTGAGTGGAGATGAGCTGCGGAGTCTGCGCTTGGCTGCGCAGGAGCCAGG 3990
QY 1587 GCTGAGTGGAGATGAGCTGCGGAGTCTGCGCTTGGCTGCGCAGGAGCCAGG 1640

RESULT 13

ID US-09-042-460-1 STANDARD; DNA; UNC; 3496 BP.

AC xxxxxx

DT

DE Sequence 1, Application US/09042460

CC Sequence 1, Application US/09042460

CC GENERAL INFORMATION:

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Allsopp, Richard

CC APPLICANT: Depinho, Ronald

CC APPLICANT: Greenberg, Roger

CC TITLE OF INVENTION: Mouse Telomerase Reverse Transcriptase

CC NUMBER OF SEQUENCES: 101

CC CORRESPONDENCE ADDRESS:

CC ADDRESS: Townsend and Townsend and Crew LLP

CC CITY: San Francisco

CC STATE: California

CC COUNTRY: USA

CC ZIP: 94111-3834

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent In Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/09/042,460

CC FILING DATE: 16-MAR-1998

CC CLASSIFICATION: 536

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/724,643

CC FILING DATE: 01-OCT-1996

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/844,419

CC FILING DATE: 18-APR-1997

CC PRIOR APPLICATION DATA:

CC

CC

CC

CC

CC

CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/974,549
CC FILING DATE: 19-NOV-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/974,584
CC FILING DATE: 19-NOV-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/979,742
CC FILING DATE: 26-NOV-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Einhorn, Gregory P.
CC REGISTRATION NUMBER: 38,440
CC REFERENCE/DOCKET NUMBER: 015389-003110US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 3496 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: -
CC LOCATION: 1..3496
CC OTHER INFORMATION: /note= "mouse telomerase reverse
CC OTHER INFORMATION: transcriptase (mTRT) cDNA clone"
CC FEATURE:
CC NAME/KEY: misc_feature
CC LOCATION: 10..3435
CC OTHER INFORMATION: /note= "mouse telomerase reverse
CC OTHER INFORMATION: transcriptase (mTRT) cDNA"
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 39..3404
CC OTHER INFORMATION: /product= "mouse telomerase reverse
CC OTHER INFORMATION: transcriptase (mTRT)"
CC SEQUENCE 3496 BP; 762 A; 998 C; 928 G; 808 T; 0 OTHER.

Query Match

Best Local Similarity 23.6%; Score 453; DB 23; Length 3496;

Matches 1216; Conservative 0; Mismatches 592; Indels 84; Gaps 3;

Db 23 CCCGGCCTTGAGCACAAATGACCCGCGCTCTCGTTGCGCCCGCGGTGCGCTCTCTGCTGCG 82

QY 51 CCCGGCACCCCGCGATGCGCGCGCTCCCGCGTCCCGCGTCCGAGCGCTGCTGCTGCG 110

Db 83 CAGCCGATACGGGAGGTGTGGCGCTGGCAACCTTTGTGGCGCGCTGGGCCCCGAGGG 142

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 111 CAGCCACTACCGCAGAGTGTCTGCCCTGGCCACAGTTCTGTGGGGCCCTGGGGCCCCAGGG 170
Db 143 CAGGGGGCTTGTGAACCCGGGGACCCGAGATCTACCGCACTTGTGTGGTGGCCCAATGCTT 202
QY 171 CTGGCGGCTGTGAGCGCGGGGACCCCGCGGCTTTCCCGCGGTGTGTGGCCCAAGTCT 230
Db 203 AGTGTGATGACACTGGGGCTCACAGCCTCCACCTGCCAGCTTCTTCCATCCACAGGTGTC 262
QY 231 GGTGTGCGTGTGGGACGACGCGCCCGCCCGCCCGCCCTCTCTTCCGCGAGGTGTC 290
Db 263 ATCCCTGAAGAGCTGTGTGGCAGGTTGTGCAGAGACTCTCGAGCGCAACAGAGAGAAA 322
QY 291 CTGCTGAAGAGCTGTGTGGCCGAGTCTGCAGAGGCTGTGCGAGCGGGCGGCGAAGAA 350
Db 323 CGTGTGCTTTTGGCTTTGAGCTGTCTTAACAGAGCCAGAGCGGGCTTCCCATGGCCCTT 382
QY 351 CGTGTGCTTTGCGCTTTCGCGCTGTGAGCGGGGCGCGGGGCGCCCGCCCGAGGCGCTT 410
Db 383 CACTAGTAGCTGGCTAGCTACTTGGCCCAACACTGTTATTGAGACCTTGGCTGTCACTGG 442
QY 411 CACCACCAAGCTGGCAGCTACTTGGCCCAACAGCGGTGACCGACACTGCGGGGGAGCGG 470
Db 443 TGCATGATGCTACTTGTAGCCGAGTGGGGACGACCTGCTGGTCTACCTGTGGGACA 502
QY 471 GCGTGGGGCTGTGCTGCGCGGTGGCGACGACGTGCTGTACCTGTCTGGCAG 530
Db 503 CTGTGCTTTTATCTTTGCTGGCCCCCAGCTGTGCTTACCAGTGTGTGGTCTTCCCT 562
QY 531 CTGCGCGCTTTTGTGTGTGGTTCCTCAGCTGTGCGCTTACCAGTGTGCGGGCGCGCT 590
Db 563 GTACCAATTTGTGCCACAGGATATCTGGCCCTCTGTGTCGCTAGTTACAGGCCAC 622
QY 591 GTACAGCTGGCGGTGCACTAGCGCGCGCCCGCCGACACGCTAGTGGACCCCGAAG 650
Db 623 CCGACCCCTGGCGAGGAATTTCACTTAACCTTAGGTTCTTTACAAGATCAAGAGCAGTAG 682
QY 651 GCGTCTGGATGCG-----AACGGGCTTGAACCATAG 683
Db 683 TCGCAGGAAGACCGGAACCCCTGGCTTGCCTATCTCGAGGTACAAAGAGCATCTGAG 742
QY 684 CGTCAGGAGGCGGGGTCTCCCTGGGCTGCGAGCCCGCGGTGCGAGGAGCGCGGG 743
Db 743 TCTCAGGATACAGTGTGCTTACGTAAAGAGCCAGATGCTATCTGTCCCGAGAGT 802
QY 744 CAGTGCCACGGAAGTCTGCGCTTGTGCCAAGAGCGCCAGGCTGTGCGCTGCGCCCTGAGCC 803
Db 803 GGAGGAGGACCCACAGCAGGCTCTACCAACCCCATCAGSCAAATCATGGTGGCCCAAG 862
QY 804 GGAGCGGAGCCCGTGGCGAGGGTCTGTGGCCCAACCGCGGAGAGCGGTGGACCCAG 863
Db 863 TCCTGTCTCGTCCCGAGGTGCTTACTGTACGAGAGAAAGATTGTCTTAAAGAAAGGT 922
QY 864 TGACCGTGTGTTCTGTGTGGTGTACCTGTCCAGAGCCCGCCGGAAGACCACTCTTTGGA 923
Db 923 GTCTGACCTGAGTCTCTTGG-----TCGTTGTGTGTAAACAGACAGCCCGAG 970
QY 924 GGGTGTGCTCTGTGACGCGCCACTCCACCCATCCGTGGGCGCCGACAGCAGCGCGG 983
Db 971 CTCCACATCTGTGCTTACCAACCCCGCCAAATGSCCTTTTCAGCTCAGGCCCATTTATGA 1030
QY 984 CCCCCATCCAGTCGCGGCCACACAGTCTCTGGACAGCGCTTGTCCCCCGGTGTAGCG 1043
Db 1031 GACCAAGACATTTCTTTACTCCAGGGGAGATGGCCAGAGGCTTTAAACCCCTCATTTCT 1090
QY 1044 CGAGACCAAGCACTTCTCTACTCTCAGCGCAACAGGAGCTGTGCGGCCCTCTCTTCT 1103
Db 1091 ACTCAGAACCTCCAGCTTACTTGTGAGTGGGGCCAGAGAGCTGTGTGAGATCATCTTTCT 1150
QY 1104 ACTCAGTCTCTGTAGGCCCGGCTGACTGTGGGCTGTGGAGGCTGTGGAGACCATCTTTCT 1163
Db 1151 GGGCTCAGGCTAGGACATCAGGACCACTCTGCAAGGACACACCTCTATCGCGTCGATA 1210
QY 1164 GGGTTCCAGGCCCTGGATGCGCAGGACTTCCCGCGAGGTGTCGCCCGCTTGGCCCGAGCGTA 1223

Db 1211 CTGCGAGATGGGGCCCTGTGTTTCCAAACAGCTGCTGGTGAACCATCGAGAGTGCCCAATATGT- 1270
QY 1224 CTGCAAAATGGGCCCTGTGTTCTGGAGCTGCTTGGAAACACCGCGAGTGCCCTACGG 1283
Db 1271 CAGACTCTCAGGTCACATTGCGAGTTTTCGAACAGCAAAACCAAGGTGACAGATGCTT 1330
QY 1284 GGTGCTCTCAAGACGACATGCGCGCTGCGAGTGCCTACCCAGCAGCGGTGTCTG 1343
Db 1331-----GAACACACCGCCACCC 1345
QY 1344 TGCCCGGAGAGCCCGAGGGCTCTGTGGCGCCCGCGAGGAGGAGACACAGACCCCG 1403
Db 1346 GCACCTCATGATTTGCTCGCCTGCACAGAGTCCCTTGGCAGGTATATGTTTCTTTCG 1405
QY 1404 TCGCTGTGTGAGTGTCTCGCCACAGACAGACGACGCGCTTGGCAGGTGTACGGCTTCGTGCG 1463
Db 1406 GGCCTGTCTCTGAAGTGGTGTCTGTAGTCTCTGGGTACCGAGGCACATGAGCGCG 1465
QY 1464 GGCCTGTCTGCGCGGTGGTGGCCCGAGGCTCTGGGGCTCCAGGCACACAGAACCGCG 1523
Db 1466 CTTCTTTAAGAACTTAAGAGTTTCATCTGTTGGGAAATACGCAAGCTATCACTGCA 1525
QY 1524 CTTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCA 1583
Db 1526 GGAATGATGTGAAGATGAAAGTAGAGGATGCGCACTGGCTCCGACAGCCCGGGGAA 1585
QY 1584 GGAGCTGACGTGGAAGATGAGCGTGGCGGAGTGGCTTGGCTGGCAGAGCCCGAGGGT 1643
Db 1586 GGACCGTGTCCCGCTGCAGAGCACCTCTGAGGAGAGGATCTGGCTAGGTTCCTGTT 1645
QY 1644 TGGCTGTGTTCCGCGCCAGAGCACCTCTGCGTGAAGGATCTGCGCAAGTTCTCTGCA 1703
Db 1646 CTGCTGATGACACATACGTGTGACAGCTGCTTAGGTGCTTTTACATCAGCAGAG 1705
QY 1704 CTGCTGATGAGTGTACGTCTGCGAGCTGCTCAGGTCTTCTTTATGTACGGAGAC 1763
Db 1706 CACATTCAGAGAACAGGCTTCTTCTACCGTAAAGAGTGTGTGGAGCAAGCTGCGAG 1765
QY 1764 CACGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAAG 1823
Db 1766 CATTGGAGTCAAGCAACACTTGAGAGAGTGGCGCTACGGAGCTGTCAAGAGGAGGT 1825
QY 1824 CATTGGAATCAGACAGCACTTGAAGAGGTGCACTCGGGAGGCTGTGCGAAGCAGAGGT 1883
Db 1826 CAGCATCACCAGCACACCTGGCTAGCCATGC 1857
QY 1884 CAGCAGATCGGAAGCCAGGCCCGCCCTGC 1915

RESULT 14

ID US-09-026-981-36 STANDARD; DNA; UNC; 3346 BP.
AC xxxxx

DT Sequence 36, Application US/09026981
CC Sequence 36, Application US/09026981
CC GENERAL INFORMATION:

CC APPLICANT: Counter, Christopher M.

CC APPLICANT: Meyerson, Matthew

CC APPLICANT: Weinberg, Robert A.

CC TITLE OF INVENTION: Telomerase Catalytic Subunit Gene and

CC NUMBER OF SEQUENCES: 52

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

CC City: Lexington

CC STATE: MA

CC COUNTRY: USA

CC ZIP: 02173

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/026,981
CC FILING DATE: 20-FEB-1998
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/064,322
CC FILING DATE: 30-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/055,762
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/054,549
CC FILING DATE: 01-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/047,151
CC FILING DATE: 20-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/038,750
CC FILING DATE: 20-FEB-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WH197-11p4AM
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 781-861-6240
CC TELEFAX: 781-861-9540
CC INFORMATION FOR SEQ ID NO: 36:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 3346 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 3346 BP; 617 A; 1027 C; 965 G; 735 T; 2 OTHER.

Query Match 14.6%; Score 281; DB 23; Length 3346;
Best Local Similarity 100.0%; Pred. No. 2,85e-234;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 971 GGGTTGGCTGTCTCGCGCCGAGACACCGTCTCGTGAGGAGATCTCTGGCCAAAGTTCC 1030
|||
Qy 1640 GGGTTGGCTGTCTCGCGCCGAGACACCGTCTCGTGAGGAGATCTCTGGCCAAAGTTCC 1699
|||
Db 1031 TGCACCTGGCTGATGAGTGACGTCGTCGAGCTGCTCAGGTCTTTCTTTTATGTCACGG 1090
|||
Qy 1700 TGCACCTGGCTGATGAGTGACGTCGTCGAGCTGCTCAGGTCTTTCTTTTATGTCACGG 1759
|||
Db 1091 AGACACCGTTTCAAAAGACAGCGCTTTTCTACCGGAGAGTCTCTGGACCAAGTTCC 1150
|||
Qy 1760 AGACACCGTTTCAAAAGACAGCGCTTTTCTACCGGAGAGTCTCTGGACCAAGTTGC 1819
|||
Db 1151 AAAGCATTGGAATCAGACAGCACTTGAAGAGGTGTCAGCTCGGGAGCTGTCTGGAAGCAG 1210
|||
Qy 1820 AAAGCATTGGAATCAGACAGCACTTGAAGAGGTGTCAGCTCGGGAGCTGTCTGGAAGCAG 1879
|||
Db 1211 AGGTCAGGAGCATCGGGAAGCCAGCGCCCGCTCTGACG 1251
|||
Qy 1880 AGGTCAGGAGCATCGGGAAGCCAGCGCCCGCTCTGACG 1920
|||

RESULT 15
ID US-08-911-312-52 STANDARD; DNA; UNC; 535 BP.
AC xxxxxx
DT

DE Sequence 52, Application US/08911312
CC Sequence 52, Application US/08911312
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.

CC APPLICANT: Andrews, William
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase
CC NUMBER OF SEQUENCES: 170
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/911,312
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Einhorn, Gregory P.
CC REGISTRATION NUMBER: 38,440
CC REFERENCE/DOCKET NUMBER: 015389-002500US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 52:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 535 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 535 BP; 105 A; 152 C; 173 G; 105 T; 0 OTHER.

Query Match 13.5%; Score 259; DB 22; Length 535;
Best Local Similarity 100.0%; Pred. No. 6.37e-213;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 AGAGCACCGTCTCGTGAGGAGATCTCTGGCCAAAGTTCTCTGACACGCTGATGAGTGTA 60
|||
Qy 1662 AGAGCACCGTCTCGTGAGGAGATCTCTGGCCAAAGTTCTCTGACACGCTGATGAGTGTA 1721
|||
Db 61 CGTGTGAGCTGCTCAGGTCTTTCTTTATGTCACGAGACCAAGTTTCAAAAGACAG 120
|||
Qy 1722 CGTGTGAGCTGCTCAGGTCTTTCTTTATGTCACGAGACCAAGTTTCAAAAGACAG 1781
|||
Db 121 GCTCTTTTCTACCGAAGAGTGTCTGAGCAAGTTGCAAGCATTTGAATCAGACAGCA 180
|||
Qy 1782 GCTCTTTTCTACCGAAGAGTGTCTGAGCAAGTTGCAAGCATTTGAATCAGACAGCA 1841
|||
Db 181 CTTGAAGAGGCTGACGCTGCGGGAGCTGTCTGGAAGACAGAGGTCAGGACATCGGGAAGC 240
|||

| Qy | 1842 | CTTTAAGAGGGTTCAGCTCGGAGGCTCTCGAAGCAGAGTCTCAGGAGCATCGGAAGC | 1901 |
|----|------|---|------|
| Db | 241 | CAGGCCCGCCTGCTGACG | 259 |
| Qy | 1902 | CAGGCCCGCCTGCTGACG | 1920 |

Search completed: Thu Dec 24 16:09:24 1998
Job time : 3547 secs.

THIS PAGE BLANK (USPTO)

W P S R E L I (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 15:07:06 1998; MasPar time 156.09 Seconds
Tabular output not generated. 573.790 Million cell updates/sec

Title: >US-08-951-733-13

Description: (1-1920) from US08951733.seq (1 of 2)

N.A. Sequence: 1920 1 CACCGGTCGGGCGAGCGCTG.....CCAGGCGCGCTGCTGACG 1920

Comp: GTGCGCAGGCGGCTGCGGAC.....GGTCCGGCGGCGAGCTGC

Scoring table: TABLE default

Gap open 30; Gap extend 1

Mismatch STD : Dbase 0; Query 0

Searched: 88822 seqs, 23323279 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: n-issued

1:5_COMB 2:PCT9_COMB 3:backfiles1

Statistics: Mean 8.977; Variance 5.692; scale 1.577

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|-----------------------------------|-----------|
| 1 | 67 | 3.5 | 7218 | 1 | US-08-232- Sequence 14, Applicati | 1.19e-24 |
| 2 | 44 | 2.3 | 7218 | 1 | US-08-232- Sequence 14, Applicati | 2.87e-11 |
| 3 | 34 | 1.8 | 215 | 1 | US-08-238- Sequence 5, Applicatio | 7.29e-06 |
| 4 | 35 | 1.8 | 215 | 1 | US-08-238- Sequence 5, Applicatio | 2.19e-06 |
| 5 | 26 | 1.4 | 74 | 2 | PCT-US95-1 Sequence 100, Applicat | 7.10e-02 |
| 6 | 26 | 1.4 | 74 | 2 | PCT-US95-1 Sequence 94, Applicati | 7.10e-02 |
| 7 | 26 | 1.4 | 81 | 2 | PCT-US95-1 Sequence 92, Applicati | 7.10e-02 |
| 8 | 26 | 1.4 | 81 | 2 | PCT-US95-1 Sequence 98, Applicati | 7.10e-02 |
| 9 | 26 | 1.4 | 82 | 2 | PCT-US95-1 Sequence 97, Applicati | 7.10e-02 |
| 10 | 26 | 1.4 | 82 | 2 | PCT-US95-1 Sequence 97, Applicati | 7.10e-02 |
| 11 | 25 | 1.3 | 66 | 1 | US-08-471- Sequence 144, Applicat | 2.09e-01 |
| 12 | 25 | 1.3 | 68 | 1 | US-07-977- Sequence 243, Applicat | 2.09e-01 |
| 13 | 25 | 1.3 | 69 | 1 | US-08-471- Sequence 142, Applicat | 2.09e-01 |
| 14 | 25 | 1.3 | 74 | 2 | PCT-US95-1 Sequence 94, Applicati | 2.09e-01 |
| 15 | 25 | 1.3 | 74 | 2 | PCT-US95-1 Sequence 100, Applicat | 2.09e-01 |
| 16 | 25 | 1.3 | 75 | 2 | PCT-US95-1 Sequence 99, Applicati | 2.09e-01 |
| 17 | 25 | 1.3 | 75 | 2 | PCT-US95-1 Sequence 99, Applicati | 2.09e-01 |
| 18 | 25 | 1.3 | 81 | 2 | PCT-US95-1 Sequence 98, Applicati | 2.09e-01 |
| 19 | 25 | 1.3 | 81 | 2 | PCT-US95-1 Sequence 92, Applicati | 2.09e-01 |

| | | | | | | | |
|---|----|----|-----|------|---|-----------------------------------|----------|
| c | 20 | 25 | 1.3 | 242 | 1 | US-08-273- Sequence 1, Applicatio | 2.09e-01 |
| c | 21 | 25 | 1.3 | 1004 | 2 | PCT-US95-0 Sequence 7, Applicatio | 2.09e-01 |
| c | 22 | 25 | 1.3 | 1386 | 2 | PCT-US95-0 Sequence 2, Applicatio | 2.09e-01 |
| c | 23 | 25 | 1.3 | 1611 | 2 | PCT-US93-0 Sequence 3, Applicatio | 2.09e-01 |
| c | 24 | 23 | 1.2 | 59 | 2 | PCT-US95-1 Sequence 95, Applicati | 1.70e+00 |
| c | 25 | 24 | 1.2 | 65 | 1 | US-08-471- Sequence 145, Applicat | 6.02e-01 |
| c | 26 | 24 | 1.2 | 66 | 2 | PCT-US95-1 Sequence 93, Applicati | 6.02e-01 |
| c | 27 | 23 | 1.2 | 66 | 1 | US-08-471- Sequence 144, Applicat | 1.70e+00 |
| c | 28 | 24 | 1.2 | 68 | 1 | US-08-471- Sequence 143, Applicat | 6.02e-01 |
| c | 29 | 23 | 1.2 | 69 | 1 | US-08-471- Sequence 142, Applicat | 1.70e+00 |
| c | 30 | 23 | 1.2 | 84 | 1 | US-08-133- Sequence 120, Applicat | 1.70e+00 |
| c | 31 | 23 | 1.2 | 84 | 1 | US-08-209- Sequence 23, Applicati | 1.70e+00 |
| c | 32 | 23 | 1.2 | 84 | 2 | PCT-US94-0 Sequence 25, Applicati | 1.70e+00 |
| c | 33 | 23 | 1.2 | 84 | 1 | US-08-300- Sequence 25, Applicati | 1.70e+00 |
| c | 34 | 23 | 1.2 | 225 | 1 | US-07-807- Sequence 15, Applicati | 1.70e+00 |
| c | 35 | 23 | 1.2 | 225 | 1 | US-08-299- Sequence 18, Applicati | 1.70e+00 |
| c | 36 | 23 | 1.2 | 1404 | 3 | 5480796-8 Patent No. 5480796 | 1.70e+00 |
| c | 37 | 23 | 1.2 | 1404 | 3 | 5171840-8 Patent No. 5171840 | 1.70e+00 |
| c | 38 | 23 | 1.2 | 1640 | 1 | US-07-807- Sequence 11, Applicati | 1.70e+00 |
| c | 39 | 23 | 1.2 | 1640 | 1 | US-08-299- Sequence 11, Applicati | 1.70e+00 |
| c | 40 | 23 | 1.2 | 2061 | 3 | 5480796-1 Patent No. 5480796 | 1.70e+00 |
| c | 41 | 23 | 1.2 | 2061 | 3 | 5171840-1 Patent No. 5171840 | 1.70e+00 |
| c | 42 | 23 | 1.2 | 2253 | 3 | 5457037-2 Patent No. 5457037 | 6.02e-01 |
| c | 43 | 24 | 1.2 | 2454 | 1 | US-07-872- Sequence 1, Applicatio | 1.70e+00 |
| c | 44 | 23 | 1.2 | 2625 | 3 | 5457037-4 Patent No. 5457037 | 1.70e+00 |
| c | 45 | 23 | 1.2 | 3336 | 3 | 5457037-1 Patent No. 5457037 | 1.70e+00 |

ALIGNMENTS

RESULT 1
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC xxxxxx
DT
DE Sequence 14, Application US/08232463
CC Sequence 14, Application US/08232463
CC Patent No. 5670367
CC GENERAL INFORMATION:
CC APPLICANT: DORNER, F.
CC APPLICANT: SCHEIFLINGER, F.
CC APPLICANT: FALKNER, F. G.
CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22313-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/232.463
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/07/935.313
CC FILING DATE:
CC APPLICATION NUMBER: EP 91 114 300.6
CC FILING DATE: 26-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703)836-9300
CC TELEFAX: (703)683-4109
CC TELEX: 899149
CC INFORMATION FOR SEQ ID NO: 14:

CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 100:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 74 BP; 6 A; 6 C; 1 G; 1 T; 60 OTHER.

Query Match 1.4%; Score 26; DB 2; Length 74;
Best Local Similarity 10.3%; Pred. No. 7.10e-02;
Matches 7; Conservative 20; Mismatches 41; Indels 0; Gaps 0;

Db 3 AGAVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 62
Cp 541 AGAGCGCGCAGCGTCCACGAGTGAACGACGCGTCTGCGCCACGCGCGCGCAGCA 482
Db 63 VNNACAC 70
Cp 481 GCCCCAC 474

RESULT 6
ID PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DT
DE Sequence 94, Application PC/TUS9511934
CC Sequence 94, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 94:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 74 BP; 3 A; 4 C; 3 G; 1 T; 63 OTHER.

Query Match 1.4%; Score 26; DB 2; Length 74;
Best Local Similarity 11.4%; Pred. No. 7.10e-02;
Matches 8; Conservative 19; Mismatches 43; Indels 0; Gaps 0;

Db 3 GAGNNBN 62
Cp 540 GAGCGCGCAGCGTCCACGAGTGAACGACGCGTCTGCGCCACGCGCGCAGCAG 481
Db 63 BNNBACGCC 72
Cp 480 CCCCACGCC 471

RESULT 7
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DT
DE Sequence 92, Application PC/TUS9511934
CC Sequence 92, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 92:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 81 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 81 BP; 3 A; 5 C; 6 G; 4 T; 63 OTHER.

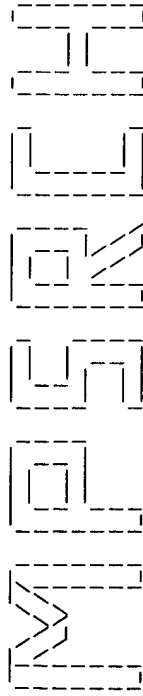
Query Match 1.4%; Score 26; DB 2; Length 81;
Best Local Similarity 11.4%; Pred. No. 7.10e-02;
Matches 8; Conservative 19; Mismatches 43; Indels 0; Gaps 0;

Db 10 GAGNNBN 69
Cp 540 GAGCGCGCAGCGTCCACGAGTGAACGACGCGTCTGCGCCACGCGCGCAGCAG 481
Db 70 BNNBACGCC 79
Cp 480 CCCCACGCC 471

RESULT 8
ID PCT-US95-11934-98 STANDARD; DNA; UNC; 81 BP.


```
AC xxxxxx
DE Sequence 98, Application PC/TUS9511934
CC Sequence 98, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 98:
CC LENGTH: 81 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 81 BP; 6 A; 6 C; 4 G; 5 T; 60 OTHER.
SQ
Query Match 1.4%; Score 26; DB 2; Length 81;
Best Local Similarity 14.3%; Pred. No. 7.10e-02;
Matches 10; Conservative 19; Mismatches 41; Indels 0; Gaps 0;
Db 1 GTTGCTCTAGANNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 60
Cp 1515 GTTGCTCTAGAGCCCGCCAGAGCCCTGGGGCCACGAGCCGCCGCGCAGCA 1456
Db 61 VNNVNNVNNV 70
Cp 1455 GCCGTACACC 1446
RESULT 9
ID PCT-US95-11934-97 STANDARD; DNA; UNC; 82 BP.
AC xxxxxx
DE Sequence 97, Application PC/TUS9511934
CC Sequence 97, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC
```

```
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 97:
CC LENGTH: 82 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 82 BP; 1 A; 2 C; 10 G; 8 T; 61 OTHER.
SQ
Query Match 1.4%; Score 26; DB 2; Length 82;
Best Local Similarity 15.2%; Pred. No. 7.10e-02;
Matches 12; Conservative 20; Mismatches 47; Indels 0; Gaps 0;
Db 2 TGTCTCTCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 61
Cp 158 TGGGGCCCCAGGGCTGGCGCTGTGTGTCAGCGCGGCGGCGGCGCTTCCGCGCGCTGG 217
Db 62 NBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBN 80
Cp 218 TGGCCCACTGCGCTGTGTG 236
RESULT 10
ID PCT-US95-11934-97 STANDARD; DNA; UNC; 82 BP.
AC xxxxxx
DE Sequence 97, Application PC/TUS9511934
CC Sequence 97, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC
```



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties
Run on: Thu Dec 24 16:09:43 1998; MasPar time 2754.39 Seconds
720.892 Million cell updates/sec
Tabular output not generated.

Title: >US-08-951-733-13
Description: (1920-2820) from US08951733.seq (2 of 2)
Perfect Score: 901
N.A. Sequence: 1920 GTCCAGACTCCGCTTCATCC.....AGCCCTGGTGGCAGCGT 2820
Comp: CAGGTCTGAGCGCAAGTAGG.....TCCGGGACCGACCGTGCCGA

Scoring table: TABLE default
Gap open 30; Gap extend 1
Nmatch STD : Dbase 0; Query 0

Searched: 567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb155
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_om 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_vl
genbank107
15:gb_ba1 16:gb_ba2 17:gb_htg 18:gb_in 19:gb_om 20:gb_ov
21:gb_pat 22:gb_ph 23:gb_pl1 24:gb_pl2 25:gb_pr1
26:gb_pr2 27:gb_pr3 28:gb_ro 29:gb_st 30:gb_sts 31:gb_sy
32:gb_un 33:gb_vl

Statistics: Mean 10.599; Variance 5.047; scale 2.100

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|---------------------------------|-----------|
| 1 | 901 | 100.0 | 4015 | 27 | AF015950 Homo sapiens telomeras | 0.00e+00 |
| 2 | 901 | 100.0 | 4027 | 27 | AF018167 Homo sapiens telomeras | 0.00e+00 |
| 3 | 901 | 100.0 | 8960 | 31 | AF043739 Synthetic construct hu | 0.00e+00 |
| 4 | 369 | 41.0 | 3369 | 28 | AF073311 Mus musculus telomeras | 7.56e-290 |
| 5 | 369 | 41.0 | 3426 | 28 | AF051911 Mus musculus telomeras | 7.56e-290 |
| 6 | 39 | 4.3 | 7218 | 21 | I66494 Sequence 14 from paten | 1.71e-09 |
| 7 | 34 | 3.8 | 74371 | 26 | AC005369 Homo sapiens chromosom | 3.54e-06 |
| 8 | 30 | 3.3 | 215 | 21 | I28278 Sequence 5 from patent | 1.16e-03 |
| 9 | 30 | 3.3 | 74371 | 26 | AC005369 Homo sapiens chromosom | 1.16e-03 |
| 10 | 29 | 3.2 | 15807 | 15 | MEU72662 Methylobacterium extor | 4.69e-03 |
| 11 | 28 | 3.1 | 215 | 21 | I28278 Sequence 5 from patent | 1.85e-02 |
| 12 | 26 | 2.9 | 5591 | 15 | AF027499 Azotobacter vinelandii | 2.66e-01 |

| | | | | | | | |
|------|----|-----|--------|----|------------|------------------------|----------|
| C 13 | 25 | 2.8 | 1908 | 28 | RRU02553 | Rattus rattus protein | 9.67e-01 |
| C 14 | 25 | 2.8 | 1912 | 28 | S81478 | CH13/CL100 Ppase-ox | 9.67e-01 |
| C 15 | 25 | 2.8 | 1948 | 28 | RNRNADSP | R.norvegicus mRNA for | 9.67e-01 |
| C 16 | 25 | 2.8 | 7218 | 21 | I66494 | Sequence 14 from paten | 9.67e-01 |
| C 17 | 24 | 2.7 | 1690 | 26 | HSRNUA14 | H.sapiens mRNA for Uli | 3.40e+00 |
| C 18 | 24 | 2.7 | 1789 | 28 | RNHLSHR | R.norvegicus HLSHR mRN | 3.40e+00 |
| C 19 | 24 | 2.7 | 2140 | 15 | S68984 | eryM-malonyl-coenzyme | 3.40e+00 |
| C 20 | 24 | 2.7 | 2140 | 15 | SERVALCOAD | Saccharopolyspora eryt | 3.40e+00 |
| C 21 | 24 | 2.7 | 2699 | 26 | AB006713 | Homo sapiens mRNA for | 3.40e+00 |
| C 22 | 24 | 2.7 | 6462 | 23 | AF044920 | Volvox carteri acetola | 3.40e+00 |
| C 23 | 24 | 2.7 | 7011 | 28 | RATRBB1 | Rat calcium channel al | 3.40e+00 |
| C 24 | 24 | 2.7 | 9695 | 28 | AF055477 | Rattus norvegicus pore | 3.40e+00 |
| C 25 | 24 | 2.7 | 26816 | 25 | HUMLUCAL16 | Human cosmid LUCAL6, c | 3.40e+00 |
| C 26 | 24 | 2.7 | 76199 | 15 | AF040570 | Amycolatopsis mediterr | 3.40e+00 |
| C 27 | 24 | 2.7 | 92879 | 17 | AC004450 | *** SEQUENCING IN PROG | 3.40e+00 |
| C 28 | 24 | 2.7 | 125532 | 17 | AC004975 | *** SEQUENCING IN PROG | 3.40e+00 |
| C 29 | 24 | 2.7 | 189370 | 15 | AF010496 | Rhodobacter capsulatus | 3.40e+00 |
| C 30 | 24 | 2.7 | 216021 | 26 | HUAC004787 | Homo sapiens Chromosom | 3.40e+00 |
| C 31 | 24 | 2.7 | 317311 | 17 | AC004085 | *** SEQUENCING IN PROG | 3.40e+00 |
| C 32 | 23 | 2.6 | 66 | 21 | I41364 | Sequence 144 from pate | 1.16e+01 |
| C 33 | 23 | 2.6 | 69 | 21 | I41362 | Sequence 142 from pate | 1.16e+01 |
| C 34 | 23 | 2.6 | 84 | 21 | I65426 | Sequence 25 from paten | 1.16e+01 |
| C 35 | 23 | 2.6 | 84 | 21 | I61313 | Sequence 120 from pate | 1.16e+01 |
| C 36 | 23 | 2.6 | 1160 | 15 | CGAJ4934 | Corynebacterium glutam | 1.16e+01 |
| C 37 | 23 | 2.6 | 2281 | 28 | MUSHAL | Mouse histidase locus | 1.16e+01 |
| C 38 | 23 | 2.6 | 4984 | 15 | AF051692 | Pseudomonas aeruginosa | 1.16e+01 |
| C 39 | 23 | 2.6 | 5636 | 18 | AB002270 | Bombyx mori L1Bm DNA, | 1.16e+01 |
| C 40 | 23 | 2.6 | 7176 | 33 | HPEGENSA | Hepatitis E virus ORF1 | 1.16e+01 |
| C 41 | 23 | 2.6 | 7200 | 33 | HPCGENOM | Hepatitais E virus com | 1.16e+01 |
| C 42 | 23 | 2.6 | 38582 | 18 | CELK12HA | C. elegans cosmid K12H | 1.16e+01 |
| C 43 | 23 | 2.6 | 53784 | 15 | AMW223012 | Amycolatopsis mediterr | 1.16e+01 |
| C 44 | 23 | 2.6 | 76199 | 15 | AF040570 | Amycolatopsis mediterr | 1.16e+01 |
| C 45 | 23 | 2.6 | 216021 | 26 | HUAC004787 | Homo sapiens Chromosom | 1.16e+01 |

ALIGNMENTS

| | | | | | |
|------------|--|--------------------|-----------------------|----------------|-------------|
| RESULT 1 | AF015950 | 4015 bp | mRNA | PRI | 16-AUG-1997 |
| LOCUS | Homo sapiens | telomerase reverse | transcriptase (hTERT) | mRNA, complete | |
| DEFINITION | cds. | | | | |
| ACCESSION | AF015950. | | | | |
| NID | 92330016 | | | | |
| KEYWORDS | human. | | | | |
| SOURCE | Homo sapiens | | | | |
| ORGANISM | Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. | | | | |
| REFERENCE | 1 (bases 1 to 4015) | | | | |
| AUTHORS | Nakamura,T.M., Morin,G.B., Chapman,K.B., Weinrich,S.L., Andrews,W.H., Lingner,J., Harley,C.B. and Cech,T.R. | | | | |
| TITLE | Telomerase catalytic subunit homologs from fission yeast and human | | | | |
| JOURNAL | Science 277 (5328), 955-959 (1997) | | | | |
| MEDLINE | 97400623 | | | | |
| REFERENCE | 2 (bases 1 to 4015) | | | | |
| AUTHORS | Morin,G.B. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (24-JUL-1997) Geron Corporation, 230 Constitution Drive, Menlo Park, CA 94025, USA | | | | |
| FEATURES | Location/Qualifiers | | | | |
| source | 1..4015 | | | | |
| gene | /organism="Homo sapiens" | | | | |
| CDS | /db_xref="taxon:9606" | | | | |
| | /tissue_type="kidney" | | | | |
| | /map="5p" | | | | |
| | /dev_stage="embryo" | | | | |
| | /chromosome="5" | | | | |
| | 1..4015 | | | | |
| | /gene="hTERT" | | | | |
| | 56..3454 | | | | |
| | /gene="hTERT" | | | | |
| | /codon_start=1 | | | | |

TKMSVRCGAWLRSPGCVPAAEHRLREEILAKFLHMLSVVYVVELLRSPFYVTF
TFOKNRLFFYKSWKSLGIRQHLKRVOLRELSEAEVROHREARPAALLTSRLRFI
PKDGLRPIVNDYVGAFTREKRAERLTSVKALFVNLTERARPGLLGASVIG
LDDIHRAMRTVLURQAQDPPPELIFVKVDVTGAYDIPQDRLTEVIAISLKPNYIC
VRYAVVQAAHGHVKAFAKSHVSTLTDQPMYRQFVAHQETSPLEDAVVIIPQSSS
NEASSGLFVDFRMCHHVRIRKGSVOCQIPQSGSILSLCSCYGMENKLFAG
IRRDGLLLRVDFLLVTHLTHAKTFLRTLVRGVEYGVNLRKTVVNFVDEAL
GSTAFVQMPAHGFLPWCGLLDDPTLEVDSDSYASARTSIRASLTNRGFKAGNRWR
KLFGLVRLKCHSLFELDLQVNSLQVCTNIYKILLQNYRHFACVQLDPHQQVWKAPT
FFURVJSDTASLSILKAKNAGSISGAGKAGALPSEAVQWICHQAFLLKLRLHRTV
YVPLGSLRTAOTLRKLPGLTTLTALEAANPALPSPDFKLTILD

BASE COUNT 674 a 1361 c 1277 g 715 t

Query Match 100.0%; Score 901; DB 27; Length 4027;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1910 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGCTGCGCGGATGTGGAACATGGACTA 1969
QY 1920 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGCTGCGCGGATGTGGAACATGGACTA 1979
Db 1970 CGTCTGGGAGCCAGAACGTTCCGACAGAAAGAGGCGGAGCGTCTCACCTCGAGGCT 2029
QY 1980 CGTCTGGGAGCCAGAACGTTCCGACAGAAAGAGGCGGAGCGTCTCACCTCGAGGCT 2039
Db 2030 GAAGGCACTGTTAGCGTCTCACTACGAGCGGCGGCGCGCGCGCTCTCTGGGCGC 2089
QY 2040 GAAGGCACTGTTAGCGTCTCACTACGAGCGGCGGCGCGCGCGCTCTCTGGGCGC 2099
Db 2090 CTCTGTGCTGGGCTGGAGCATATCCAGAGGCGTGGCGACCTTCGTGCTGGCTGGCG 2149
QY 2100 CTCTGTGCTGGGCTGGAGCATATCCAGAGGCGTGGCGACCTTCGTGCTGGCTGGCG 2159
Db 2150 GGCCAGGACCCGCGCTGAGCTGTACTTTGTCAAGTGTGATGACGGGCGGCTACGA 2209
QY 2160 GGCCAGGACCCGCGCTGAGCTGTACTTTGTCAAGTGTGATGACGGGCGGCTACGA 2219
Db 2210 CACATCCCCAGGACAGGCTCAGGAGGTGATCGGACGATATCAAAACCCAGAACAC 2269
QY 2220 CACATCCCCAGGACAGGCTCAGGAGGTGATCGGACGATATCAAAACCCAGAACAC 2279
Db 2270 GTACTGGTGGTGGTATGCGGTGCTCCAGAGGCGCGCCATGGGACGTCGCCAAGGC 2329
QY 2280 GTACTGGTGGTGGTATGCGGTGCTCCAGAGGCGCGCCATGGGACGTCGCCAAGGC 2339
Db 2330 CTTCAAGAGCCACGCTCTACTTTGACAGACCTCCAGCGGTACATGACAGTTCGTGGC 2389
QY 2340 CTTCAAGAGCCACGCTCTACTTTGACAGACCTCCAGCGGTACATGACAGTTCGTGGC 2399
Db 2390 TCACCTGAGGAGACACGCGCTGAGGATGCGGTGCTATCGAGCAGAGCTCCTCCCT 2449
QY 2400 TCACCTGAGGAGACACGCGCTGAGGATGCGGTGCTATCGAGCAGAGCTCCTCCCT 2459
Db 2450 GAATGAGCCAGCAGTGGCTTTCGAGCTTCTTCTAGCTTTCATGTCGACACGCGGT 2509
QY 2460 GAATGAGCCAGCAGTGGCTTTCGAGCTTCTTCTAGCTTTCATGTCGACACGCGGT 2519
Db 2510 GCGCATCAGGGGCAAGTCTCTACGTCAGTGCAGGGGATCCCGAGGCTCCATCTCTC 2569
QY 2520 GCGCATCAGGGGCAAGTCTCTACGTCAGTGCAGGGGATCCCGAGGCTCCATCTCTC 2579
Db 2570 CACGCTGCTCAGCCTGTGCTACGCGGACATGGAGAACAGCTGTTTGGCGGATTCG 2629
QY 2580 CACGCTGCTCAGCCTGTGCTACGCGGACATGGAGAACAGCTGTTTGGCGGATTCG 2639
Db 2630 GCGGACGGGCTGCTCTGCGCTTTGGTGGATGATTTCTTGTGTGTGACACCTCACCTCAC 2689
QY 2640 GCGGACGGGCTGCTCTGCGCTTTGGTGGATGATTTCTTGTGTGTGACACCTCACCTCAC 2699
Db 2690 CCACGCAAAACCTTCTCTCAGSACCTGGTCCGAGGCTGCTCCTGATGATGCTCGGTGT 2749
QY 2700 CCACGCAAAACCTTCTCTCAGSACCTGGTCCGAGGCTGCTCCTGATGATGCTCGGTGT 2759

Db 2750 GAACCTCCGGAACACAGCTGTGACTTCCCTGTAGAGACGAGGCGCTGGTGCCACGGC 2809
QY 2760 GAACCTCCGGAACACAGCTGTGACTTCCCTGTAGAGACGAGGCGCTGGTGCCACGGC 2819
Db 2810 T 2810
QY 2820 T 2820
RESULT 3
LOCUS AF043739 8960 bp DNA SVN 05-APR-1998
DEFINITION Synthetic construct human telomerase catalytic subunit
(hTERT/hEST2) mRNA and 3' Influenza A hemagglutinin tag, complete
sequence.
ACCESSION AF043739
NID G3023054
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 8960)
AUTHORS Counter,C.M., Meyerson,M., Eaton,E.N., Ellisen,L.W., Caddle,S.D.,
Haber,D.A. and Weinberg,R.A.
TITLE Telomerase activity is restored in human cells by ectopic
expression of hTERT (hEST2), the catalytic subunit of telomerase
Oncogene 16 (9), 1217-1222 (1998)
JOURNAL 9818713
MEDLINE
REFERENCE 2 (bases 1 to 8960)
AUTHORS Counter,C.M., Meyerson,M., Eaton,E.N., Ellisen,L.W., Caddle,S.D.,
Haber,D. and Weinberg,R.A.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1998) Whitehead Institute, 9 Cambridge Center,
Cambridge, MA 02142, USA
FEATURES
Location/Qualifiers
1..8960
/organism="synthetic construct"
/db_xref="taxon:32830"
order(1..1101,4606..8960)
/note="mammalian expression vector; see GenBank Accession
Number U47120"
/organism="Cloning vector pCI-neo"
/db_xref="taxon:45772"
1102..4542
/note="telomerase catalytic subunit hTERT/hEST2 mRNA"
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5p15.33"
1102..4542
/gene="hTERT/hEST2"
1147..4605
/note="telomerase/reverse transcriptase"
/codon_start=1
/transl_table=11
/product="telomerase catalytic subunit"
/db_xref="PID:G3023055"
/translation="MPRAPRCRAVRSLRSHYREVLPLATFVRRLPGQGNRLVQRGDP
AARALVAOCLVCPNDARPPPAAPSFQVSCLELYVARVLQRLCERGAKNVLAFOGA
LIDGARGGPEAFTTVSYRLPNTVDALRGSGAWGLLLLRVGDVLLVLLARCALFV
LVAPSCAYQVCGPPLYGATQAPPPHAGSPRRRLGCRANWNSVREAGVPLGPA
PGARRGGASRLPLKRPARGAAPERTPVGQSWAHPTGSRGTRGPGSCFVSP
RPAEATSLGSLGTRHSGPSRQHGAGPPSTSRPPRWDTPCPVVAETKHFLYS
SGDEOLRPSFLLSLPSITGARRLVETFLGSRWPMCTPRRLPQRYWQMRPL
FLELLNHAQCPYGVLLKTHCPRLAAVTAAGVCAREKPGQSVAAAEEDTDPRLLVQ
LLHQSSPMQVYGFVRACLRLLVPPGLMGRNRHNRFLRNTKFIISLGHAKVLSLQEL
TKMSVRCGAWLRSPGCVPAAEHRLREEILAKFLHMLSVVYVVELLRSPFYVTF
TFOKNRLFFYKSWKSLGIRQHLKRVOLRELSEAEVROHREARPAALLTSRLRFI
PKDGLRPIVNDYVGAFTREKRAERLTSVKALFVNLTERARPGLLGASVIG
LDDIHRAMRTVLURQAQDPPPELIFVKVDVTGAYDIPQDRLTEVIAISLKPNYIC
VRYAVVQAAHGHVKAFAKSHVSTLTDQPMYRQFVAHQETSPLEDAVVIIPQSSS
NEASSGLFVDFRMCHHVRIRKGSVOCQIPQSGSILSLCSCYGMENKLFAG
IRRDGLLLRVDFLLVTHLTHAKTFLRTLVRGVEYGVNLRKTVVNFVDEAL
GSTAFVQMPAHGFLPWCGLLDDPTLEVDSDSYASARTSIRASLTNRGFKAGNRWR
KLFGLVRLKCHSLFELDLQVNSLQVCTNIYKILLQNYRHFACVQLDPHQQVWKAPT
FFURVJSDTASLSILKAKNAGSISGAGKAGALPSEAVQWICHQAFLLKLRLHRTV
YVPLGSLRTAOTLRKLPGLTTLTALEAANPALPSPDFKLTILD

GGTAFAVMPAHGLFPWCGLLDTRTFLEQSDYSYARTSIRASITFNRGFKAGNMR
KLEGLRLKCHSLFDLQVNSLOVCTNIVKILLQAYREHACVLOLPFFHOVKNPT
FFLRVSDTSLCYSLKAKNAGMSLGAAGPLSEAVOMLCHQAFLLKLRHRT
YPLUGSLRATQOLSKRLPGLTTLTALEAANPALPSDFKLTILDYDVPDYAGPYD
VPDYAG"

source

4543. .4605

/note="engineered based on the epitope for the
anti-influenza virus hemagglutinin monoclonal antibody
12CA5; similar to part of the Influenza A hemagglutinin
gene sequence in GenBank Accession Number M29257"
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 1886 a 2551 c 2484 g 2038 t 1 others
ORIGIN

Query Match 100.0%; Score 901; DB 31; Length 8960;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;

Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3000 GTCCAGACTCGCTCATCCCAAGCCGTGACGGGCTGGCGGATGTGAACATGGACTA 3059
|||||
Qy 1920 GTCCAGACTCGCTCATCCCAAGCCGTGACGGGCTGGCGGATGTGAACATGGACTA 1979
|||||
Db 3060 GTCGTGGGAGCCAGAGCTTCCGAGAGAAAGAGGCGCGGCTCTCACTCGAGGCT 3119
|||||
Qy 1980 GTCGTGGGAGCCAGAGCTTCCGAGAGAAAGAGGCGCGGCTCTCACTCGAGGCT 2039
|||||
Db 3120 GAAGCACTGTCAGCGTCTCACTAGAGCGGCGGCGGCGGCGGCTCTCGGGCGC 3179
|||||
Qy 2040 GAAGCACTGTCAGCGTCTCACTAGAGCGGCGGCGGCGGCGGCTCTCGGGCGC 2099
|||||
Db 3180 CTCGTGCTGGGCGTGAGAGATATCCAGAGGCGCTGGCGACCTTCGTGCTGCTGCG 3239
|||||
Qy 2100 CTCGTGCTGGGCGTGAGAGATATCCAGAGGCGCTGGCGACCTTCGTGCTGCTGCG 2159
|||||
Db 3240 GCGCCAGGACCGCGCTGAGCTTACTTGTCAAGTGGATGACGGCGGCTAGCA 3299
|||||
Qy 2160 GCGCCAGGACCGCGCTGAGCTTACTTGTCAAGTGGATGACGGCGGCTAGCA 2219
|||||
Db 3300 CACCATCCCCAGGACAGGCTCAGGAGGTCTATCGCCAGCATCATCAACCCCAAGAAC 3359
|||||
Qy 2220 CACCATCCCCAGGACAGGCTCAGGAGGTCTATCGCCAGCATCATCAACCCCAAGAAC 2279
|||||
Db 3360 GTACTGCGTGGTGGTATGCGGTGTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGG 3419
|||||
Qy 2280 GTACTGCGTGGTGGTATGCGGTGTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGG 2339
|||||
Db 3420 CTTCAAGAGCCAGCTCTACTTGCACAGACTCCAGCGCTACATGCGCAGAGTTCGTGGC 3479
|||||
Qy 2340 CTTCAAGAGCCAGCTCTACTTGCACAGACTCCAGCGCTACATGCGCAGAGTTCGTGGC 2399
|||||
Db 3480 TCACCTGAGGAGACACCGCGCTGAGGATGCCGTGCTCATCGAGCAGAGCTCTCCCT 3539
|||||
Qy 2400 TCACCTGAGGAGACACCGCGCTGAGGATGCCGTGCTCATCGAGCAGAGCTCTCCCT 2459
|||||
Db 3540 GAATGAGGCGCAGCTGCGCTCTTCGAGCTCTTCCTAGCTTCATGTCGACACGCGCT 3599
|||||
Qy 2460 GAATGAGGCGCAGCTGCGCTCTTCGAGCTCTTCCTAGCTTCATGTCGACACGCGCT 2519
|||||
Db 3600 GCGCATAGGCGCAGCTCTAGCTTCAGCTGCGGAGATCCGCGAGGCTCCATCTCTC 3659
|||||
Qy 2520 GCGCATAGGCGCAGCTCTAGCTTCAGCTGCGGAGATCCGCGAGGCTCCATCTCTC 2579
|||||
Db 3660 CACGTGCTCTGAGCCTCTGCTAGCGGACATGAGAGAACAGCTGTTGCGGGGATTCG 3719
|||||
Qy 2580 CACGTGCTCTGAGCCTCTGCTAGCGGACATGAGAGAACAGCTGTTGCGGGGATTCG 2639
|||||
Db 3720 GCGGACCGGCTGCTCTGCGCTTGGTGGATGATTTCTTGTGTGACACCTCAGCTCAC 3779
|||||
Qy 2640 GCGGACCGGCTGCTCTGCGCTTGGTGGATGATTTCTTGTGTGACACCTCAGCTCAC 2699
|||||
Db 3780 CCACGCAAAACCTTCCTCAGGACCTGTCGAGGCTGTCCTGAGTATGCTGCTGCTG 3839
|||||

Qy 2700 CCACGCGAAACCTCTCCTCAGGACCTCTCCTCAGGCTGTCCTCAGTATGCTCGTGGT 2759
Db 3840 GAACCTCGGAAGACACTGGTGAACCTTCCCTGTAGAGACGAGCCCTGGTGCACGCGC 3899
|||||
Qy 2760 GAACCTCGGAAGACACTGGTGAACCTTCCCTGTAGAGACGAGCCCTGGTGCACGCGC 2819
|||||
Db 3900 T 3900
Qy 2820 T 2820

RESULT 4

LOCUS AF073311 3369 bp mRNA ROD 09-SEP-1998
DEFINITION Mus musculus telomerase catalytic subunit mRNA, complete cds.
ACCESSION AF073311
NID g3551846
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 3369)
AUTHORS Martin-Rivera, L., Herrera, E., Albar, J. P. and Blasco, M. A.
TITLE Expression of mouse telomerase catalytic subunit in embryos and adult tissues

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (18), 10471-10476 (1998)

MEDLINE 98393668

REFERENCE 2 (bases 1 to 3369)

AUTHORS Martin-Rivera, L., Herrera, E. and Blasco, M. A.

TITLE Direct Submission

JOURNAL Submitted (19-JUN-1998) Immunology and Oncology, National Centre of Biotechnology, Cantoblanco, Madrid 28049, Spain

FEATURES
Location/Qualifiers
1. .3369

/organism="Mus musculus"

/db_xref="taxon:10090"

/codon_start=1

/product="telomerase catalytic subunit"

/translation="MTRAPRCAPVRSLLRSRYREVWPLATFVRLQPEGRRLVQPGDP
KIYRTVAOCLVCMHWSQPPADLSHOVSSILKELVARVORLCERNERNVLAFGFE
LNEARGGPPMAFTSSVRSYLPNTVETLRVSGAWMLLSRNGDLDLLYLALCALYL
LVPSCAYQVCGSLFQICATDINFSVSASIRTPRPVVEGHPHQVLPFTSGKSWSPR
APKALPAPRGTKRHLSLTSTSVPSAKKACYPVPRVEGHPHQVLPFTSGKSWSPR
ARSPVPTAEKDLSSKGVSDLSGKVCCKHKPSSTSLSPRQAFQLRPFTRH
FLYSGDQGERLNPSSLNLPNLGARRLVEIIFLSRPTSGPLRHLRSRYV
QMPPLFQQLLVNHAECQVYVLLPSHCRFTANQVTDALNTSPHMDLLRLHSSPQ
VYGLFACILKVVYASLWGTNRHNERFFNLKKEFISLGYKGLSLQELMWMKVECH
WLKSSPGKDRVPAAEHRLERILATFLWMDTVVOLLRSFFYITESTFQRRLFFY
RKSWSKLQSIGVROHLERVRELSEVEYRHQDWTLMAPICRLRFIPKNGLRPIV
NMSYSGTRALGRKQAHFTQRLKTLFSLNLYERTKHPHMGSSVLGMDIYRTWRA
FVLVRALDQTPRMVFKADVTGAYDAIPQGLVEVYVANNIRHSESTYCIROYAVVR
DSOGVYKSPRQVTTLSLOPYMGOFKHLQSDASALNSVVIQSIISMNSSL
FDFLHFLRHSVVKIGDKYTCQGIPOGSSSLTLCSCFCGDMENKLFAYEVDGLL
LRVDDFLVTTPLHDQKFTSLTVHGVPEYGCMLNKTVMNFPVPEPTGLGAAPYQ
LKCHGLFPMWGLDQTLVEFCDYSYQATSIKTSITFSQFKAQKRNKLLSVLR
LPAGHFLDLOVNSLQVTCINIKYIFLQAYRHFACVIQIOPDQVRKNLTFFLGIS
SQASCCVAILKVKNPMTLKASGFPPEAAHWLCYQAFLLKLAHSVIYKCLLGLPLT
AQKLCKRLPEATMTILKAADPALSTDFQILD"

BASE COUNT 737 a 959 c 891 g 782 t

ORIGIN

Query Match 41.0%; Score 369; DB 28; Length 3369;

Best Local Similarity 73.5%; Pred. No. 7.56e-290;

Matches 667; Conservative 0; Mismatches 231; Indels 9; Gaps 2;

Db 1827 CAGACTCGGCTTATCCCAAGCCCAAGCGCTCGCGCCCATGTGACATFAGTTATAG 1886
|||||

Qy 1923 CAGACTCGGCTTATCCCAAGCCCAAGCGCTCGCGCCCATGTGACATFAGTTATAG 1982
|||||

Db 1887 CATGGTACACAGACTTTGGCGCAGAGGACGAGCCGACGATTTACCCAGCGCTCAA 1946
|||||

QY 1983 CGTGGGACGACAGCTTCCGACAGAAAGAGCGCGCTCACCCTCGAGGTGAA 2042
Db 1947 GACTCTCTCAGCATGCTCAACTATGAGCGGCAAAACATCTCCTACCTTAAGGGTCTTC 2006
QY 2043 GGCAGCTGCTCAGGCTGCTCAACTACGAGCGGCGCGCGCTCCTCGGCGCTC 2102
Db 2007 TGACTGGGTATGATGACATCTACAGGACCTGCGGGGCTTGTGCTGGTGGCTGC 2066
QY 2103 TGTCTGGGCTGACCATATCACAGGGCTGCGGCACCTGCTGCTGGTGGCGGC 2162
Db 2067 TCTGGACGACACCCAGGATGCTTGTAGGCGAGATGTGACCGGGCTATGATGC 2126
QY 2163 CCAGGACCCGCGCTCAGCTGCTACTTGTCAAGGTGATGTGACGGCGCTACGACAC 2222
Db 2127 CATCCCCAGGGTAAGCTGTGGAGTGTGTGCAATATGATAGGCACTCGGAGACAC 2186
QY 2223 CATCCCCAGGACAGGCTCAGGAGGTGATCGGCGAGCATCATCAACCC---AGAACAC 2279
Db 2187 GTACTGTATCGGACGATGACAGTGTGCGGAGAGATAGCCAGGCCAAGTCCACAGTC 2246
QY 2280 GTACTGGCTGCGGTATGCGGTGGTCCAGAGGCGCGCCCATGGGACCGTCCGCAAGC 2339
Db 2247 CTTTAGGACAGCTCACCACCTCTCTGACCTCCAGCATATCATGCGGCGCTTCTCTAA 2306
QY 2340 CTTCAAGAGCCAGCTCTCTACCTTGACAGACCTCCAGCGCTACATGCGAGTCTGTGC 2399
Db 2307 GCATCTCAGGATTCAGATGCCAGTGCAGTGCAGGAACTCCGTGTGATCGAGCAGACAT 2366
QY 2400 TCACCTCAGAGAC-----CAGCCCGCTGAGGATGCGGTCTCATCGAGCAGCTC 2453
Db 2367 CTCTAAGTATGAGAGCAGCAGCAGCTGTTGACTTCTTCTGCACTTCTGCTGCTACAG 2426
QY 2454 CTCCCTGAATGAGCCAGCAGTGGCTCTTCCAGCGTCTTCTAGCTTCTATGTGCCACCA 2513
Db 2427 TGTCTAAGATTTGTCAGAGTCTATACAGCTGCAGGCGCATCCCGAGGCTCCAG 2486
QY 2514 CGCGTCCGATCAGGGCGAGTCTTACGTTTGTGATGACTTCTGTTGAGCGCTCA 2606
QY 2634 GATTCGGCGGAGCGGCTGCTCTGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2693
Db 2607 CTTGACCAAGCAAAACCTTCTCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2666
QY 2694 CTTACCCACCGAAGAACCTTCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2753
Db 2667 CATGATAACTTGCAGAGACAGTGGTGAACCTTCCCTGTGAGGCTGCTGCTGCTGCTGCTG 2726
QY 2754 CGTGGTGAATTCGGAAGACAGTGGTGAACCTTCCCTGTAGAGCAGGCGCTGCTGCTGCTG 2813
Db 2727 TGCAGCT 2733
QY 2814 CACGGCT 2820

RESULT 5
LOCUS AF051911 3426 bp mRNA ROD 02-APR-1998
DEFINITION Mus musculus telomerase reverse transcriptase mRNA, complete cds.
ACCESSION AF051911
NID g3005591
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3426)
AUTHORS Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.
TITLE Expression of mouse telomerase reverse transcriptase during development, differentiation, and proliferation

JOURNAL Oncogene (1998) In press
REFERENCE 2 (bases 1 to 3426)
AUTHORS Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-1998) Microbiology and Immunology, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA
FEATURES
source Location/Qualifiers
1. 3426
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="13"
/map="MGD offset position 40; 40 cM from centromere"
30. 3398
/note="mTERT"
/codon_start=1
/product="telomerase reverse transcriptase"
/db_xref="pid:g3005592"
/translation="MTRAPRCFAVRLSRVREVPVLFATFVRRLGPEGRRLVQPDPE
KIYRLVAQCLVCMHWGSPADLSFHQVSSKLKELVARVQRLKERNRNVLAFGE
LNEARGPPMAFTSVRSYLPNTVETLRVSGAWMLLSRVGDDLLVYLAHCALYL
LVPPCAYQVCGSPLYQICATDIWPSVASYPTRPVGRNTNRLFLQIKSSSROE
APKPLALPSGRTKRLSLTSTVSPSAKAKCYPPVREBPHROVLPFTSGSKWSPSP
ARSPVPTAEKDLSSKGKVDLSLSGVCCCKHPSTSLSPRQNAFOLRPIETRH
FLYSRGDOERLNPFLNSQPNUTGARRLVEIIFLGSRPRTSGPLCHRLSRRIW
QMRPLFQQLLVNHAECQVYRLLRSHCRFTANQVTDALNTSPHLDMLLHSSPMQ
VYGFRLACLKVVYSASLWGRHNRERFKLKKFISLGRYKLSQLSDLMKRWKVEDCH
WURSSPKQSGIGVROHLRRLRELSQEVRRHODTLAMPICRLRFIPKPNGLRFTV
RKSYSKLSQSGVROHLRRLRELSQEVRRHODTLAMPICRLRFIPKPNGLRFTV
NMSYSGTALGRKQAQHTORLKTLSMLNYERTKHPHMGSSVLGMDNYRTWRA
FVIRVALDQTPMTFVKADVTGADYDAIPQGLVEVYVANNRHSSTYICRIQAVVRR
DSGQVHKFSFRQVTTLSDLQMGOFKHLQSDASALRNVSTEQISMNSSL
FDFFLHLSVYKIGDRCTQCGIPQGSLSLTLCLSCFDMENKMLFAEYQDGLL
LRFVDFLLVPLHQAQFTSLTVHGVPEYGMINLQKTVNPEVPTGLGAAPYQ
LPAHCLFPMCGLLDTOTLEVFCDYSGYASIKTSLTFSQSVFKAGKTWRNKLVLRL
LACHGLDLQVNSLQTVCLINIKYIFLQAYRHFACVIOLPFDQVRKNLTFLGLIIS
SQASCCYATLKVNFGMTLKSGSFPPEAAHMLCQAFLLKLAHSSVIYKCLLGPURT
AOKLLCRKLPETATILKAAADPALSTDFQTILD"

BASE COUNT 746 a 979 c 907 g 794 t

Query Match 41.0%; Score 369; DB 28; Length 3426;
Best Local Similarity 73.5%; Pred. No. 7.56e-230;
Matches 667; Conservative 0; Mismatches 291; Indels 9; Gaps 2;

Db 1856 CAGACTGCGCTTCATCCCAAGCCCAAGCGCTCGCGCCCATTTGTAACATGATATAG 1915
QY 1923 CAGACTCGGCTTCATCCCAAGCGCTGCGCGCCCATTTGTAACATGATATAG 1982
Db 1916 CATGGGTACAGAGCTTTGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1975
QY 1983 CGTGGGAGCCAGAACGTTCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2042
Db 1976 GACTCTCTCAGCATGCTCAACTATGAGCGGCAAAACATCTCCTACCTTAAGGGTCTTC 2035
QY 2043 GGCATGTTTCAGGCTGCTCAACTACGAGCGGCGCGCGCGCGCGCGCTC 2102
Db 2036 TGTACTGGGTATGAATGACATCTACAGGACCTGCGGGCGCTTGTGCTGCTGCTGCTG 2095
QY 2103 TGTGCTGGGCTGGAGATATCCAGGCGCTGCGGACCTTCGCTGCTGCTGCTGCTGCTG 2162
Db 2096 TCTGGACACAGACCCAGGATGCTTGTAAAGGAGAGATGTGACCGGGCGCTATGATGC 2155
QY 2163 CCAGGACCCCGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2222
Db 2156 CATCCCCAGGGTAAGCTGTGGAGGTGTTGTCATATATGATAGGCACTCGGAGACAC 2215
QY 2223 CATCCCCAGGACAGGCTCAGGAGGTGATCGGAGCATCATCAACCC---AGAACAC 2279
Db 2216 GTACTGATCCGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2275
QY 2280 GTACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2339


```

repeat_region 4366. 4661
/rpt_family="Alu"
repeat_region 5327. 5602
/rpt_family="Alu"
repeat_region 6586. 6956
/rpt_family="L1"
repeat_region 6647. 6684
/note="(CA)19"
/rpt_type-tandem
/rpt_unit-CA
7113. 7373
/rpt_family="Alu"
repeat_region complement(7830. 8185)
/standard_name="possible repeat"
repeat_region 8238. 8503
/rpt_family="Alu"
repeat_region 9070. 9387
/rpt_family="Alu"
repeat_region complement(9740. 9845)
/rpt_family="MER42"
repeat_region complement(10440. 11015)
/rpt_family="Alu"
repeat_region 11950. 12250
/rpt_family="Alu"
repeat_region 12657. 12085
/note="(A)29"
/rpt_type-tandem
/rpt_unit-A
repeat_region 12365. 12645
/rpt_family="Alu"
repeat_region 13727. 13750
/note="(AC)12"
/rpt_type-tandem
/rpt_unit-AC
repeat_region 13783. 14024
/rpt_family="L1"
repeat_region 14175. 14470
/rpt_family="Alu"
repeat_region complement(14906. 15259)
/standard_name="possible repeat"
repeat_region 15300. 15613
/rpt_family="Alu"
repeat_region 16671. 16690
/note="(T)20"
/rpt_type-tandem
/rpt_unit-T
repeat_region complement(16675. 16977)
/rpt_family="Alu"
repeat_region complement(16993. 17085)
/rpt_family="MER42"
repeat_region complement(17678. 18276)
/rpt_family="Alu"
repeat_region 19305. 19583
/rpt_family="Alu"
repeat_region 19914. 19945
/note="(GTTT)8"
/rpt_type-tandem
/rpt_unit-GTTT
repeat_region complement(19943. 20222)
/rpt_family="Alu"
repeat_region 21202. 21496
/rpt_family="Alu"
repeat_region 21736. 22035
/rpt_family="Alu"
repeat_region 22017. 22038
/note="(A)22"
/rpt_type-tandem
/rpt_unit-A
repeat_region 22321. 22457
/rpt_family="Alu"
repeat_region complement(22638. 22981)
/standard_name="possible repeat"
23022. 23326

```

```

repeat_region /rpt_family="Alu"
23473. 23761
/rpt_family="Alu"
repeat_region 23744. 23767
/note="(A)24"
/rpt_type-tandem
/rpt_unit-A
repeat_region complement(24639. 24694)
/rpt_family="MER42"
repeat_region complement(25349. 25713)
/rpt_family="Alu"
repeat_region complement(25727. 26471)
/rpt_family="Alu"
27191. 27477
/rpt_family="Alu"
repeat_region 27774. 28057
/rpt_family="Alu"
repeat_region 28040. 28066
/note="(A)27"
/rpt_type-tandem
/rpt_unit-A
repeat_region 28712. 28930
/rpt_family="MER20"
repeat_region complement(28769. 28838)
/note="GRAIL 2 excellent exon, frame 0"
28987. 29214
/rpt_family="Alu"
misc_feature join(29485. 29559,29598. 29779)
/note="95% identity dbSTS:G14522 (SHCC-11312)"
29495. 29976
/standard_name="AI027942"
/note="100% identity EST ov84a10.x1"
complement(30401. 30536)
/note="GRAIL 2 excellent exon, frame 2"
complement(30682. 30733)
/rpt_family="MIR"
misc_feature complement(31573. 31724)
/note="GRAIL 2 excellent exon, frame 1"
complement(32159. 32232)
/note="GRAIL 2 excellent exon, frame 2"
32388. 32488
/rpt_family="MLT1"
repeat_region 32617. 32908
/rpt_family="Alu"
repeat_region 32977. 33088
/rpt_family="MLT1"
repeat_region complement(33670. 33785)
/rpt_family="Alu"
misc_feature complement(34021. 34144)
/note="GRAIL 2 excellent exon, frame 2"
complement(35238. 35331)
/note="GRAIL 2 excellent exon, frame 1"
36392. 36663
/rpt_family="Alu"
misc_feature 36901. 37222
/note="100% identity EST ou55c09.x1"
/db_xref="dbEST:AI025011"
36901. 37164
/standard_name="STSC-9983"
/db_xref="dbSTS:G26554"
complement(37595. 37654)
/note="GRAIL 2 excellent exon, frame 0"
complement(join(38063. 38218,38462. 38578,38741. 38995,
39071. 39203,39532. 39630,39935. 40048,40300. 40410,
40503. 40661,41868. 41972,42103. 42225,42492. 42569,
44379. 44507))
/standard_name="histidyl-tRNA synthetase"
/note="78% 100% protein identity GenPept:U18937"
complement(38069. 38215)
/note="GRAIL 2 excellent exon, frame 0"
complement(38462. 38578)

```

...
Note: remainder of annotations omitted.

Query Match 3.8%; Score 34; DB 26; Length 74371;
Best Local Similarity 15.5%; Pred. NO. 3.54e-06;
Matches 13; Conservative 46; Mismatches 25; Indels 0; Gaps 0;

Db 15918 KCKCTSTKSCMSKSRGKRGYSWKYRCAMWMTCKSSKCCWCSYEMRMKCYSCSYC 15977
Cp 2609 TCGCGCTAGCAGCTGCAGACGACGCTGGAGAGATGGACCCCTGGCGGATCCCTGG 2550

Db 15978 YCSSGKKYWCRCSMYWTYYSYK 16001
Cp 2549 CACTGGACGTAGACTGCCCTG 2526

RESULT 8
LOCUS 128278 215 bp DNA PAT 30-OCT-1996
DEFINITION Sequence 5 from patent US 5569830.
ACCESSION 128278
NID 91819054
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 215)
AUTHORS Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
TITLE Plant inhibitors of fungal polygalacturonases and their use to control fungal disease
JOURNAL Patent: US 5569830-A 5 29-OCT-1996;
FEATURES Location/Qualifiers
source 1..215
/organism="unknown"

BASE COUNT 15 a 8 c 25 g 26 t 141 others
ORIGIN

Query Match 3.3%; Score 30; DB 21; Length 215;
Best Local Similarity 14.8%; Pred. NO. 1.16e-03;
Matches 16; Conservative 46; Mismatches 46; Indels 0; Gaps 0;

Db 1 MTNVMTSSSVSRVASCNDKAKDGNITSSWTTDCNRTNVCOTDTYRVNNDSGHNK 60
QY 2107 CTGGCGCTGAGATATCCACAGCGCTTGGCGCACCTTCTGCTGGTGGCGGCCAG 2166
Db 61 YSANYNNGNNVGAHYYHTHTVNSGADSKTVTDVSNASTSSNG 108
QY 2167 GACCCCGCCCTGAGCTGTACTTGTCTCAAGGTGGATGTGACGGGCGC 2214

RESULT 9
LOCUS AC005369 74371 bp DNA PRI 01-AUG-1998
DEFINITION Homo sapiens chromosome 5, BAC clone 119j3 (LBNL H175), complete sequence.
ACCESSION AC005369
NID 93367505
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 74371)
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE Sequencing of human chromosome 5
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 74371)
Ricke,D.O.
TITLE Large Scale Sequence Analysis and Annotation with the Sequence Comparison Analysis (SCAN) System
JOURNAL Unpublished
AUTHORS 3 (bases 1 to 74371)
Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.

TITLE Direct Submission
JOURNAL Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, MS 74-157, Berkeley, CA 94720, U.S.A.
COMMENT Sequence submitted by:
DOE Joint Genome Institute.
FEATURES Location/Qualifiers
source 1..74371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="5g"
/clone="119j3"
/chromosome="5"
/note="LBNL H175"
893..1030
/rpt_family="Alu"
2255..2438
/rpt_family="Alu"
2818..2859
/note="(GT)21"
/rpt_type=tandem
/rpt_unit=GT
join(3246..3410,3721..3828)
/standard_name="RUF"
/note="65% & 69% protein identity GenPept:U22377"
3431..3724
/rpt_family="Alu"
3707..3728
/note="(A)22"
/rpt_type=tandem
/rpt_unit=A
4366..4661
/rpt_family="Alu"
5327..5602
/rpt_family="Alu"
6586..6956
/rpt_family="L1"
6647..6684
/note="(CA)19"
/rpt_type=tandem
/rpt_unit=CA
7113..7373
/rpt_family="Alu"
complement(7830..8185)
/standard_name="possible repeat"
8258..8503
/rpt_family="Alu"
9070..9387
/rpt_family="Alu"
complement(9740..9845)
/rpt_family="MER42"
complement(10440..11015)
/rpt_family="Alu"
11950..12250
/rpt_family="Alu"
12057..12085
/note="(A)29"
/rpt_type=tandem
/rpt_unit=A
12365..12645
/rpt_family="Alu"
13727..13750
/note="(AC)12"
/rpt_type=tandem
/rpt_unit=AC
13783..14024
/rpt_family="L1"
14175..14470
/rpt_family="Alu"
complement(14906..15259)
/standard_name="possible repeat"
15300..15613
/rpt_family="Alu"

| | | |
|---|---------------|--|
| | repeat_region | 16671..16908 /note="(T)20" /rpt_type=tandem /rpt_init=T complement(16675..16977) /rpt_family="Alu" complement(16993..17085) /rpt_family="MER42" complement(17678..18276) /rpt_family="Alu" 19305..19583 /rpt_family="Alu" 19914..19945 /note="(GTTT)8" /rpt_type=tandem /rpt_init=GTTT complement(19943..20222) /rpt_family="Alu" 21202..21196 /rpt_family="Alu" 21736..22035 /rpt_family="Alu" 22017..22038 /note="(A)22" /rpt_type=tandem /rpt_init=A 22321..22457 /rpt_family="Alu" complement(22638..22981) /standard_name="possible repeat" 23022..23326 /rpt_family="Alu" 23473..23761 /rpt_family="Alu" 23744..23767 /note="(A)24" /rpt_type=tandem /rpt_init=A complement(24639..24694) /rpt_family="MER42" complement(25349..25713) /rpt_family="Alu" complement(25727..26471) /rpt_family="Alu" 27191..27477 /rpt_family="Alu" 27774..28057 /rpt_family="Alu" 28040..28066 /note="(A)27" /rpt_type=tandem /rpt_init=A 28712..28930 /rpt_family="MER20" complement(28769..28838) /note="GRAIL 2 excellent exon, frame 0" 28987..29214 /rpt_family="Alu" Join(29485..29559,29598..29779) /note=".95% identity dbSTS:G14522 (SHGC-11312)" 29495..29976 /standard_name="AT027942" /note="100% identity EST ov84a10.x1" complement(30401..30536) /note="GRAIL 2 excellent exon, frame 2" complement(30682..30733) /rpt_family="MIR" complement(31573..31724) /note="GRAIL 2 excellent exon, frame 1" complement(32159..32232) /note="GRAIL 2 excellent exon, frame 2" 32388..32488 /rpt_family="MLT1" |
| | repeat_region | 32617..32908 /rpt_family="Alu" 32977..33088 /rpt_family="MLT1" complement(33670..33785) /rpt_family="Alu" complement(34021..34144) /note="GRAIL 2 excellent exon, frame 2" complement(35238..35331) /note="GRAIL 2 excellent exon, frame 1" 36392..36663 /rpt_family="Alu" 36901..37222 /note="100% identity EST ou55c09.x1" /db_xref="dbEST:AI025011" 36901..37164 /standard_name="STSG-9983" /db_xref="dbSTS:G26554" complement(37595..37634) /note="GRAIL 2 excellent exon, frame 0" 39071..39205,39532..39630,39935..40048,40300..40410, 40503..40661,41868..41972,42103..42225,42492..42569, 44379..44507)) /standard_name="histidyl-tRNA synthetase" /note="78%-100% protein identity GenPept:U18937" complement(38069..38215) /note="GRAIL 2 excellent exon, frame 0" complement(38462..38578) |
| | misc_feature | Query Match 3.3%; Score 30; DB 26; Length 74371; Best Local Similarity 14.7%; Pred. No. 1.16e-03; Matches 20; Conservative 63; Mismatches 53; Indels 0; Gaps 0; |
| D | b | 15906 WGAASRGSWSTCKCTKCTSTSKSCMSRSKRGRGWYRKWKRCAMMTCKSSKCWCWSY 15965 ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: Qy 2042 AGGCACTGTTCAGCGTGCTCAACTACGAGCGGGCCGCCGCCCTCCTGGCGCCT 2101 |
| D | b | 15966 RMRMKCYCSVCYCISGKKYWCRSMYYTYCYYSKYKYKSNSVYTCTSWGRWNWSKGSR 16025 : : ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: |
| Q | y | 2102 CTGTGCTGGCGCTGGACGATATCACAGGGCCTGGCGCACCTTCGTGCTGTGCGGG 2161 :: :: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: |
| D | b | 16026 WMVASRGSCSCSMC 16041 :: :: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: |
| Q | y | 2162 CCRAGGACC CGCC 2177 :: :: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: |
| R | E | SULT 10 DNA BCT 06-JUN-1997 LOCUS MEU72662 15807 bp DEFINITION Methylobacterium extorquens methylophily region containing methyl-CoA lyase (mclA), putative ABC transporter subunit A (abcA), putative ABC transporter subunit B (abcb), putative ABC transporter ATP-binding subunit (abcc), methanol dehydrogenase large subunit homolog (mxar'), cytochrome c (maxg'), mxaj homolog (mxaj'), 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (folA), dihydronopterin aldolase (folB) and dihydroxyterate synthase (folC) genes, complete cds, and phosphoenolpyruvate carboxylase (ppca) gene, partial cds, and pyrroquinoline quinone biosynthesis gene cluster containing PqqE (pqqE) gene, complete cds, and PqqC/D gene, partial cds. U72662 U60993 gi577783 |
| A | C | CESSION NID KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS TITLE JOURNAL |
| M | e | thylobacterium extorquens. Methylobacterium extorquens Eubacteria; Proteobacteria; alpha subdivision; Methylobacterium. 1 (bases 1 to 13697) Chistoserdova,L and Lidstrom,M.E. Molecular and mutational analysis of a DNA region separating two methylophily gene clusters in Methylobacterium extorquens AM1 Microbiology 143 (Pt 5), 1729-1736 (1997) |


```

/codon_start=1
/transl_table=11
/product="alginase"
/db_xref="PID:g2920567"
/translacion="MKTRLALSCLLGSLLLSGAVHAAEALVPPKGVYAPVDIRKGEA
PACVPVEPTGLVFRSKVSGSDAARSTLNEAEAKFRKTAPIQTQIERGVSEWYMR
YMEKGRAGDLECTLAWDAWAEDGALLTEYNTGSKMRKWAUGSLAGAYLRUKFTSS
QPLAAYPEQRIHAGQVDSNFKVDQVQIKDSDLPKRNNHSYWAAMVAAAGVATNRRP
LFDAWYEQFHIAAGVDSNFKVPELNRQRALAYNHSYSLPPLMVAAMAFALANGVDLR
FDWAGALGRAGNLAGVKEPFAERAGDEDDMEDLETDAKESWLEPYCALYSCSP
ALPRRAEMGPPFRNRLGGDVTRIEDPAEKSPRSTVGKRD"
gene
2897..4396
/genes="algi"
/genes="algi"
/genes="algi"
/genes="algi"
/notes="similar to Pseudomonas aeruginosa AlgI"
/codon_start=1
/transl_table=11
/product="Algi"
/db_xref="PID:g2920568"
/translacion="WVSSNVFLFLPLFLGLYLSPARYNLLLLTASYVYAWVR
IDFLGFAAVTVNVLGRIGRAGVTSQARWLTLGVVVDLCVLGYKYANFGVDS
LNAIITSGVEPVVTHLLPGISFVVPESISYIIDVYEGDTPATRNVLVDAFAFV
FPHLIAGPVLRFRLVDQFNHRHTLDKFAEGCTREMQGFKKKVTSLSUAPADHCFR
ALSDPITDAGALAYTAOLYDFDSGIDMAGLGMGFRMFENFNQYISQSITE
FWRWHISLSTWRDLYKITSLSGNRGTTFTYRNRLITMLLGLLGHGANWTFIIGAW
HGTWLAIERALRDAAPKTIIRPLRVKFFVLMVGVVIFRAENLDVAVRWRYAAMFSG
DWTLSDLNKAQLTSLOIATLLAVVIAYVGIQFVAQPLTGAPKAKANDOADAPQI
VHASSGTLAYSOAIDIPALATRVAVLLLFAASVLKLSAQSFSPFLYQF"
gene
4409..5575
/genes="algv"
/genes="algv"
/genes="algv"
/genes="algv"
/notes="similar to Pseudomonas aeruginosa AlgJ"
/codon_start=1
/transl_table=11
/product="AlgJ"
/db_xref="PID:g2920569"
/translacion="MNRNTLVAGTTFAGTLLALSLWSLGAAGFSTADNTPYLNGKL
ALAFEHYDEFFPIKLTGNLVAALDYTLFEGRGVVICANOWLPSDEEKKTAAS
QNTIDNQALIQGVRELRANNOVLWAILPAKARLPENFGEQQPASHLELLCNFRF
IVADAGIQADPLIGLQQAAGGVQMLTHTWTPQAQVVAQOLATIKPIGVLPES
GVYVETLPGGPKGDLTNFLDPLFELLPDQAKHNTQQQFESAPAGDDLFA
ETQVVALVGTYSADERNWFAGALQALGSDLVNFAEDGRGLLPMLKFLQSEDFK
SPRLVWEPPERYLPMAYDLSEFDADWIAQLKAAGRQDLADNTATNQGRH"
BASE COUNT 1051 a 2011 c 1640 g 889 t
ORIGIN

Query Match 2.9%; Score 26; DB 15; Length 5591;
Best Local Similarity 82.5%; Pred. No. 2.66e-01;
Matches 33; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 3830 ATCTCTACCATGCTGCTCGGGGGCTGTGGACACGGCGCCA 3869
||||| ||| ||||| || ||||| || ||||| || ||||| ||
Qy 2572 ATCTCTCCAGCTGCTCTGCAGCGCTGTGTACTACGGCGACA 2611

RESULT 13
LOCUS RR02553 1908 bp DNA ROD 19-OCT-1993
DEFINITION Rattus rattus protein tyrosine phosphatase gene, complete cds.
ACCESSION U02553
NID g409022
KEYWORDS black rat.
SOURCE Rattus rattus
ORGANISM Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Rattus.
REFERENCE 1 (bases 1 to 1905)
AUTHORS Qian, Z., Gilbert, M. E. and Kandel, E. R.
TITLE Differential induction by neuronal activity in rat brain of a
protein tyrosine phosphatase containing a nuclear-localization
signal

```

Unpublished
 2 (bases 1 to 1908)
 Qian, Z.
 Direct Submission
 Submitted (15-OCT-1993) Zhuo Qian, Center For Neurobiology and
 Behaviour, Howard Hughes Medical Institute, Columbia University,
 722 West 168th Street, New York, NY 10032 USA
 Location/Qualifiers
 1. .1908
 /organism="Rattus rattus"
 /db_xref="taxon:10117"
 /dev_stage="adult"
 /tissue_type="brain"
 127. .1230
 /codon_start=1
 /product="protein tyrosine phosphatase"
 /db_xref="PID:g409023"
 /translation="WMVEGILDAGGLRALLRERRAAOCLLLDCRSFFAFNAGHIVGSV
 NYRFTSVIRRRAGKAMQVTFVQLGGYEAFASCPCLSCPMGLSLPSTSVDPDSAESG
 ALAAGALCRRRAGTQVFFLGLSAYHASRDKMLDGLTALINSAKPNHFEGHY
 CSCSTPLVDGCPVEILLFLYLSGAYHASRDKMLDGLTALINSAKPNHFEGHY
 QYKSIPEVDHNHRAIDFSIKDAGGRVFEVHCQAGISRSATICLAYLMRT
 NRKVLDEAFEFYKQRSIIISPNFSWGQLQFESQVLAHPCSAEAGSPAMAVLDGRGTS
 TITVFNPVSPVPHPTNSALNTLQSPITTPSC"

polyA_signal
 1886. .1891

BASE COUNT 408 a 540 c 496 g 464 t

ORIGIN

Query Match 2.88; Score 25; DB 28; Length 1908;
 Best Local Similarity 71.28; Pred. No. 9.67e-01;
 Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Db 142 GGCATCTGACGCGGGGGGCTGCGCGCGCTGCTGCGAGAGCGCGCGCTCAGTGCCT 200
 |||||
 Cp 2100 GCGCCAGAGGCGGGGGCGCGCGCCCTGCTAGTGCAGCAGCTGACACAGTGCCT 2042

RESULT 14
 LOCUS S81478 1912 bp mRNA ROD 29-MAY-1996
 DEFINITION 3CH134/CL100 PTPase=oxidative stress-inducible protein tyrosine
 phosphatase [rats, peritoneal macrophage cDNA library, mRNA, 1912
 nt].
 ACCESSION S81478
 NID g1336815
 KEYWORDS Rattus sp. peritoneal macrophage cDNA library.
 SOURCE Rattus sp.
 Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 1912)
 Feng, L., Xia, Y., Seifert, D. and Wilson, C.B.
 Oxidative stress-inducible protein tyrosine phosphatase in
 glomerulonephritis
 JOURNAL Kidney Int. 48 (6), 1920-1928 (1995)
 MEDLINE 96163253
 REMARK GenBank staff at the National Library of Medicine created this
 entry [NCBI g1336815] from the original journal article.
 This sequence comes from Fig. 1.
 Location/Qualifiers
 1. .1912
 /organism="Rattus sp."
 /db_xref="taxon:10118"
 115. .1218
 /gene="3CH134/CL100 PTPase"
 115. .1218
 /gene="3CH134/CL100 PTPase"
 /note="oxidative stress-inducible protein tyrosine
 phosphatase; This sequence comes from Fig. 1"
 /codon_start=1
 /product="3CH134/CL100 PTPase"
 /db_xref="PID:g1336816"
 /translation="WMVEGILDAGGLRALLRERRAAOCLLLDCRSFFAFNAGHIVGSV"

NVRFSTIVRRKAGMGLBHIVPNTELRGLLAGAYHVVLLDRSAALOGAKRDGTL
 ALAAGALCRERSTQVFSIQGGEAFSASPCELCSKTPMGLSLPSTSPVDSAS
 CSSCTPLVDQGGPGEFSLVLYGAYSHASCKRMDLALGTALLNVANGCPNFEHG
 QYKSIYPVEDNHRADISSWEALDIDSIKAAGGEVFVHCQAGISRSATTCLAYLWRT
 NVRLDEAFEFKVPKNSALNPSFMGOLLQFESQVLAPHCSADGAGSPNAVLDRTGS
 TTTVFVNFVHPTNSALNPSMGLTSTPSC"

| BASE COUNT | 426 a | 536 c | 488 g | 462 t |
|------------|-------|-------|-------|-------|
| ORIGIN | | | | |

Query Match 2.8%; Score 25; DB 28; Length 1912;
Best Local Similarity 71.2%; Pred. NO. 9.67e-01;
Matches 42; Conservative 0; Mismatches 17; Indels

Db 130 GGATCTCGACGCGGGGGGTGCGCGCGTCTGCAGAGCGGCCCTCAGTGCCCT 188
 ||| ||| ||||| ||| |||| | ||| ||| | |||||
Cp 2100 GCGCCCCAGAGCGCGGGCGCGCGCTGTTGAGCACGCTGAACAGTGCCCT 2042

| RESULT | 15 | | | | |
|------------|----|---|---------|-----|-------------|
| LOCUS | | RNRNADSP | 1948 bp | RNA | |
| DEFINITION | | R.norvegicus mRNA for dual specificity phosphate. | | | |
| | | | | ROD | 26-JAN-1995 |

NID 9642264
 KEYWORDS c100 gene; protein-tyrosine-phosphatase.
 SOURCE Norway rat.

ORGANISM Rattus norvegicus
Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Rattus.

REFERENCE 1 (bases 1 to 1948)
 AUTHORS Muda, M., Schlegel, W. and Arkinstall, S.
 TITLE Pathways regulating CL100 gene expression in pituitary cells
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1948)

TITLE Direct Submission
JOURNAL Submitted (17-JAN-1995) M. Mada, Glaxo Institute for Molecular Biology, 14 chemin des Aulx, CH-1228 Plan-les-Ouates, Geneva, SWITZERLAND

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .1948 |

```
1: 13960
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/dev_stage="adult"
/tissue_type="lung"
/clone_lib="lambda Grl0"
/clone_id="Grl100"
```

gene

CDS
/gene="cll100"
164. .1267

```
/gene="cll100"  
/EC_number="3.1.3.48"  
/note="protein-tyrosine-phosphatase"  
/codon_start=1  
/product="dual specificity phosphatase"
```

[illegible]

| POLYALSIGNAL 1924. .1929 | | | |
|--------------------------|-------|-------|-------|
| BASE COUNT | 420 a | 544 c | 515 g |
| ORIGIN | | | |
| | | | 469 t |

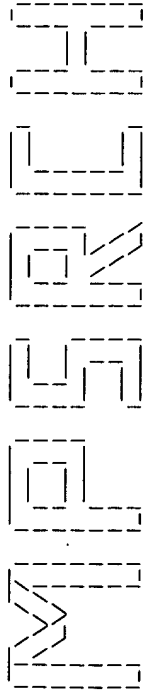
Query Match 2.8%; Score 25; DB 28; Length 1948;
Best Local Similarity 71.2%; Pred. No. 9.67e-01;
Matches 42; Conservative 0; Mismatches 17; Indels

Db 179 GGCATCCTGGACGCCGGGGCTGCGCGCTGCTGCGAGAGCGCGCTCAGTGCCT 237

Cp 2100 GGGCCCCAGGAGCGCGGGGGCGCGCCCGCTCTAGTTGAGCACGCTGAACAGTGCCT 2042.

Search completed: Thu Dec 24 17:24:07 1998
Job time : 4464 secs.

THIS PAGE BLANK (USPTO)



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 17:24:26 1998; MasPar time 2466.70 seconds
Tabular output not generated. 654.110 Million cell updates/sec

Title: >US-08-951-733-13
Description: (1920-2820) from US08951733.seq (2 of 2)
Perfect Score: 901
N.A. Sequence: 1920 GTCCAGACTCGCTTCATCC.....AGGCCCTGGGTGGCAGCGGT 2820
Comp: CAGGCTCTGAGCGAAGTAGG.....TCCGGGACCCACCGTGGCGA

Scoring table:
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-est55
Database: 1:em_est1 2:em_gss1 3:em_gss2 4:em_gss3
genbank-est107

5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est20 17:gb_est21
18:gb_est22 19:gb_est23 20:gb_est24 21:gb_est25 22:gb_est26
23:gb_est27 24:gb_est28 25:gb_est29 26:gb_gss1 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 10.621; Variance 2.221; scale 4.783

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | ID | Description | Pred. No. |
|------------|-------|-------------|----------|-------------------------|-----------|
| 1 | 161 | 17.9 | AA311750 | EST182469 Jurkat T-cell | 7.58e-249 |
| 2 | 159 | 17.6 | AA281296 | zT08g02.r1 NCI_CGAP_GC | 5.81e-245 |
| 3 | 49 | 5.4 | AA754459 | 97SN1787 Rice Immature | 7.23e-42 |
| 4 | 47 | 5.2 | AA754459 | 97SN1787 Rice Immature | 1.31e-38 |
| 5 | 46 | 5.1 | AA754458 | 97SN1784 Rice Immature | 5.39e-37 |
| 6 | 44 | 4.9 | AA754458 | 97SN1784 Rice Immature | 8.51e-34 |
| 7 | 33 | 3.7 | AA200728 | mul3h09.r1 Soares 2MBM | 4.00e-17 |
| 8 | 33 | 3.7 | AA200728 | mul3h09.r1 Soares 2MBM | 4.00e-17 |
| 9 | 26 | 2.9 | AA034173 | Homo sapiens ntcon2 co | 4.10e-07 |
| 10 | 26 | 2.9 | AA098755 | T3985 MVAT4 bloodstrea | 1.10e-07 |
| 11 | 26 | 2.9 | T26788 | T529 Trypanosoma bruce | 1.10e-07 |
| 12 | 26 | 2.9 | AA034177 | Homo sapiens ntcon6 co | 1.10e-07 |
| | | | AF038250 | Homo sapiens clone ntc | 1.10e-07 |

| | | | | | | | |
|---|----|----|-----|---------|----------|-------------------------|----------|
| C | 13 | 26 | 2.9 | 2275 11 | AF034173 | Homo sapiens ntcon2 co | 1.10e-07 |
| | 14 | 25 | 2.8 | 301 14 | AA848961 | EST191723 Normalized r | 1.91e-06 |
| | 15 | 25 | 2.8 | 318 14 | AA799774 | EST191723 Normalized r | 1.91e-06 |
| C | 16 | 25 | 2.8 | 660 11 | AF034177 | Homo sapiens ntcon6 co | 1.91e-06 |
| | 17 | 24 | 2.7 | 148 10 | AA619013 | vo6g06.r1 Soares mous | 3.08e-05 |
| | 18 | 24 | 2.7 | 299 21 | W04101 | T1794 MVAT4 bloodstrea | 3.08e-05 |
| C | 19 | 24 | 2.7 | 325 15 | A1029813 | UI-R-C0-jn-c-04-0-UI.s | 3.08e-05 |
| | 20 | 24 | 2.7 | 358 18 | A103459 | EST121748 Normalized r | 3.08e-05 |
| | 21 | 24 | 2.7 | 389 21 | W69043 | T2954 MVAT4 bloodstrea | 3.08e-05 |
| C | 22 | 24 | 2.7 | 405 24 | AA065279 | el0500r Testis 5 Homo | 3.08e-05 |
| | 23 | 24 | 2.7 | 418 15 | A1008367 | ES20202818 Normalized r | 3.08e-05 |
| C | 24 | 23 | 2.6 | 175 8 | C29002 | Rise cDNA, partial seq | 4.56e-04 |
| | 25 | 23 | 2.6 | 209 18 | A1078208 | Oy5a01.xl NCI_CGAP_CL | 4.56e-04 |
| | 26 | 23 | 2.6 | 279 17 | AA555543 | ZF-E218 zebrafish embr | 4.56e-04 |
| C | 27 | 23 | 2.6 | 312 15 | AA972885 | op24g09.s1 Soares_NFL | 4.56e-04 |
| | 28 | 23 | 2.6 | 334 11 | AA716778 | zg71f09.s1 Soares_feta | 4.56e-04 |
| C | 29 | 23 | 2.6 | 339 12 | AA760318 | vv71d10.r1 Stratagene | 4.56e-04 |
| | 30 | 23 | 2.6 | 357 24 | AA065280 | e01500c Testis 5 Homo | 4.56e-04 |
| C | 31 | 23 | 2.6 | 367 15 | AA947159 | on5a05.s1 Soares_NFL | 4.56e-04 |
| | 32 | 23 | 2.6 | 374 23 | AA166027 | ms25a01.r1 Stratagene | 4.56e-04 |
| C | 33 | 23 | 2.6 | 397 8 | AA283952 | zs46h12.s1 NCI_CGAP_GC | 4.56e-04 |
| | 34 | 23 | 2.6 | 414 19 | T26785 | T198 Trypanosoma bruce | 4.56e-04 |
| C | 35 | 23 | 2.6 | 441 11 | AA693472 | ah20e08.s1 Gessler Wil | 4.56e-04 |
| | 36 | 23 | 2.6 | 459 18 | AT083671 | oy7le10.xl NCI_CGAP_CL | 4.56e-04 |
| C | 37 | 23 | 2.6 | 470 21 | W90935 | mf83g12.r1 Soares mous | 4.56e-04 |
| | 38 | 23 | 2.6 | 471 8 | AA280884 | zs97e01.s1 NCI_CGAP_GC | 4.56e-04 |
| C | 39 | 23 | 2.6 | 524 24 | AA277163 | va82d07.r1 Soares mous | 4.56e-04 |
| | 40 | 23 | 2.6 | 551 21 | AA032626 | mi34g02.r1 Soares mous | 4.56e-04 |
| C | 41 | 23 | 2.6 | 583 18 | AU035221 | Mus musculus cDNA, clo | 4.56e-04 |
| | 42 | 23 | 2.6 | 612 24 | AA270172 | va9a10.r1 Soares mous | 4.56e-04 |
| C | 43 | 23 | 2.6 | 621 9 | AA589365 | vl46b07.s1 Stratagene | 4.56e-04 |
| | 44 | 23 | 2.6 | 628 15 | AA995794 | os29f10.s1 NCI_CGAP_K1 | 4.56e-04 |
| C | 45 | 23 | 2.6 | 823 17 | AT115886 | ue96b03.y1 Sugano mous | 4.56e-04 |

ALIGNMENTS

| | | | | | | |
|------------|---|-------------------|--------------|---------|-----------|-------------|
| RESULT | 1 | AA311750 | 409 bp | mRNA | EST | 19-APR-1997 |
| LOCUS | EST182469 | Jurkat T-cells VI | Homo sapiens | CDNA 5' | end, mRNA | |
| DEFINITION | sequence. | | | | | |
| ACCESSION | AA311750 | | | | | |
| NID | G1964077 | | | | | |
| KEYWORDS | EST. | | | | | |
| SOURCE | human. | | | | | |
| ORGANISM | Homo sapiens | | | | | |
| | Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; | | | | | |
| | Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; | | | | | |
| | Homo. | | | | | |
| REFERENCE | 1 (bases 1 to 409) | | | | | |
| AUTHORS | Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitchugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S. Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C. | | | | | |
| TITLE | Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence | | | | | |
| JOURNAL | Nature 377 (6547 Suppl), 3-174 (1995) | | | | | |
| MEDLINE | 96026280 | | | | | |
| COMMENT | | | | | | |

Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavet@igr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES

source
1. 409
/organism="Homo sapiens"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"
/db_xref="ATCC (inhost):158964"
/db_xref="taxon:9606"
/clone_lib="Jurkat T-cells VI"
/cell_type="T-lymphocyte"
<1..>409

BASE COUNT 65 a 120 c 133 g 86 t 5 others
ORIGIN
Query Match 17.9%; Score 161; DB 24; Length 409;
Best Local Similarity 100.0%; Pred. No. 7.58e-249;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GTTGGTGGATGATTCCTGTTGGTGACACCTCACCTCACCCACCGGAAACCTTCCTCA 60
QY 2660 GTTGGTGGATGATTCCTGTTGGTGACACCTCACCTCACCCACCGGAAACCTTCCTCA 2719
Db 61 GGACCTCGTCCGAGGTGCTCCTAGTATGCTGCTGGTGAACCTGCGGAGACAGTGG 120
QY 2720 GGACCTCGTCCGAGGTGCTCCTAGTATGCTGCTGGTGAACCTGCGGAGACAGTGG 2779
Db 121 TGAACCTCCCTAGACGAGGCCCTGGGTGGCAGCGCT 161
QY 2780 TGAACCTCCCTAGACGAGGCCCTGGGTGGCAGCGCT 2820

RESULT 2 AA281296 389 bp mRNA EST 14-AUG-1997
LOCUS
DEFINITION Zt08g02.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:712562 5',
mRNA sequence.
ACCESSION AA281296
NID g1924194
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Hominoidea;
Homo.
REFERENCE 1 (bases 1 to 389)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2187 Std Error: 0.00
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 385.
Location/Qualifiers

FEATURES

source
1..389
/organism="Homo sapiens"
/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),

provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTACCAATCTGAAGTGGAGCGGCGCTCATTTTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_lib="IMAGF:712562"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
<1..>389

BASE COUNT

ORIGIN 87 a 102 c 123 g 77 t
Query Match 17.6%; Score 159; DB 8; Length 389;
Best Local Similarity 100.0%; Pred. No. 5.81e-245;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 231 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGCTCGGGCGGATTGTGAACATGGACTA 290
QY 1920 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGCTCGGGCGGATTGTGAACATGGACTA 1979
Db 291 CGTCGTGGAGCCAGAACGTTCCGCAGAGAAAGAGGCGCGCTCTCACCTCGAGGGT 350
QY 1980 CGTCGTGGAGCCAGAACGTTCCGCAGAGAAAGAGGCGCGCTCTCACCTCGAGGGT 2039
Db 351 GAAGGCACTGTTACGCTGCTCAACTAGCAGCGGCGCGG 389
QY 2040 GAAGGCACTGTTACGCTGCTCAACTAGCAGCGGCGCGG 2078

RESULT 3 AA754459 252 bp mRNA EST 20-JAN-1998
LOCUS
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
cDNA clone 97SN1787, mRNA sequence.
ACCESSION AA754459
NID 92801165
KEYWORDS EST.
SOURCE rice.

ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
REFERENCE 1 (bases 1 to 252)
AUTHORS Nahn,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT

Contact: Eun M.Y.
Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyungido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sunn20.asti.re.kr
Submitted by Baek Hie Nahn, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhnam@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
Location/Qualifiers

FEATURES

source
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"

Db 79 TBYSNVNDTNGTGVGKTTVNVHSCWNNRCSNSVYVWBTAYCDYBHHBDRANHVDDT 138
 QY 2411 AGACACCGCCGCTGAGGATCGCTGTCATCGAGCAGAGCTCCTCCTGAATGAGGCCA 2470
 Db 139 RCTNDRGVCNVTASDNGTSATKRTVGDYDKDSCGGCKRWKTVGSBYBRCGVNVNVRT 198
 QY 2471 GCAGTGCCCTTCGAGCTCTTCTAGCTTCATGTGCCACCGCGCGGATCAGG 2530
 Db 199 TSMWTDKSTKMSMDMSRRSRVHVGRW 225
 QY 2531 GCAGTCTCAGTCCAGTCCAGGGA 2557

RESULT 6
 LOCUS 247 bp mRNA EST 20-JAN-1998
 DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII CDNA Library Oryza sativa
 ACCESSION AA754458
 NID AA754458
 KEYWORDS 92801164
 SOURCE EST.
 ORGANISM rice.
 Oryza sativa
 Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
 Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
 Poales; Poaceae; Oryza.
 REFERENCE 1 (bases 1 to 247)
 AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
 Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
 Lee,M.C. and Eun,M.Y.
 TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
 JOURNAL Unpublished (1998)
 COMMENT
 Contact: Eun M.Y.
 Department of Cytoogenetics
 National Inst. of Agri. Sci. and Tech, RDA
 Suwon, Kyunggi-do, Korea
 Tel: 82 331 290 0301
 Fax: 82 331 290 0307
 Email: myeunesun20.asti.re.kr
 Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
 University, Yongin, Korea. 449-728 bhnaem@bioserver.myongji.ac.kr
 Seq primer: M13 Reverse Primer.
 FEATURES
 source Location/Qualifiers
 1. .247
 /organism="Oryza sativa"
 /cultivar="Milyang23"
 /note="Vector: pBluescript SK(+); Site.1: EcoRI; Site.2:
 XhoI; Directional CDNA library inserted into lambda ZAPII
 vector at 5' end with EcoRI and 3' end with Xho I site."
 /db_xref="taxon:4530"
 /clone="97SN1784"
 /clone_lib="Rice Immature Seed Lambda ZAPII CDNA Library"
 /tissue_type="Immature Seed"
 /dev_stage="5 days after pollination"
 /lab_host="E. coli SOLR"
 BASE COUNT 7 a 16 c 21 g 34 t 169 others
 ORIGIN
 Query Match 4.9%; Score 44; DB 12; Length 247;
 Best Local Similarity 12.7%; Pred. No. 8.51e-34;
 Matches 14; Conservative 60; Mismatches 36; Indels 0; Gaps 0;
 Db 111 SNSVYVWBTAYCDYBHHBDRANHVDDTCTNDRGVCNVTASDNGTSATKRTVGYDKTDS 170
 Cp 2133 GGCCCTGTGATATCGTCAGGCCACACAGAGCCGCCAGGAGCGCGCGCGC 2074
 Db 171 DCGGCKRWKTVGSBYBRCGVNVNVRTTSMWTDKSTKMSMDMSRRSRV 220
 Cp 2073 CCGCTCTAGTTGACGACGCTGAACAGTGCCTTCACCTCGAGGTGAGAC 2024

RESULT 7
 LOCUS 2275 bp mRNA EST 22-DEC-1997
 DEFINITION Homo sapiens ntcon2 contig mRNA, partial sequence, mRNA sequence.

LOCUS AA200728 375 bp mRNA EST 19-FEB-1997
 DEFINITION mul3h09.r1 Soares 2NBMt Mus musculus cDNA clone 639329 5', mRNA
 ACCESSION AA200728
 NID q1795738
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata;
 Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Mus.
 REFERENCE 1 (bases 1 to 375)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilton,R. and
 Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:391321
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 361.
 FEATURES
 source Location/Qualifiers
 1. .375
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTTCACCAATCTGAAGTGGAGCGCGGTTTTTTTTTTTTTTTTTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through two
 rounds of normalization, and was constructed by Bento
 Soares and M.Fatima Bonaldo."
 /db_xref="taxon:10090"
 /clone="639329"
 /clone_lib="Soares 2NBMt"
 /sex="male"
 /tissue_type="Thymus"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 BASE COUNT 80 a 95 c 105 g 95 t
 ORIGIN
 Query Match 3.7%; Score 33; DB 23; Length 375;
 Best Local Similarity 75.4%; Pred. No. 4.00e-17;
 Matches 49; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 Db 144 AGAGCATCTCTATGAATGAGAGCAGCAGCAGCTGTTGACTTCTCTGCACTTCCTGC 203
 QY 2447 AGAGCTCTCTGATGAGCGCAGTCAGTCTCTTCGACGCTCTTCACGCTTCATGT 2506
 Db 204 GTCAC 208
 QY 2507 GCCAC 2511

RESULT 8
 LOCUS AF034173 2275 bp mRNA EST 22-DEC-1997
 DEFINITION Homo sapiens ntcon2 contig mRNA, partial sequence, mRNA sequence.

| | |
|--|---|
| ACCESSION | AF034173 |
| NID | 92707735 |
| KEYWORDS | EST. |
| SOURCE | human. |
| ORGANISM | Homo sapiens |
| Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; | |
| Primates; Catarrhini; Hominoidea; Homo. | |
| REFERENCE | 1 (bases 1 to 2275) |
| AUTHORS | Tripodis,N. and Ragoussis,J. |
| TITLE | Generation of a transcription map in the region immediately centromeric to human MHC across the 6p21.2-6p21.3 chromosomal boundary |
| JOURNAL | Unpublished |
| REFERENCE | 2 (bases 1 to 2275) |
| AUTHORS | Tripodis,N. and Ragoussis,J. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (13-NOV-1997) Division of Medical and Molecular Genetics, |
| FEATURES | Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK Location/Qualifiers 1..2275 |
| source | /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="6" /map="6p21.3" /clone="ntcon2 contig" /tissue_type="liver; brain" /dev_stage="fetus" /note="similar to Br140" |
| BASE COUNT | 438 a 619 c 470 g 599 t 149 others |
| ORIGIN | |
| Query Match | 3.7%; Score 33; DB 11; Length 2275; |
| Best Local Similarity | 13.3%; Pred. No. 4.00e-17; |
| Matches | 11; Conservative 47; Mismatches 25; Indels 0; Gaps 0; |
| Db | 1523 RGKKCKYKRYKYTYTKYSRWYTYTYYWCWCTSMKSASCAMRWWGYMGR 1582 |
| QY | 2648 GGTCGTCCTCGCTTGCTGGATGATTCTGTGTGGTGACACCTCACCTCACCACGCCGA 2707 |
| Db | 1583 SRSYWGYGWSGCGYGMTRYX 1605 |
| QY | 2708 AAACCTTCCTCAGGACCCGTGC 2730 |
| RESULT | 9 |
| LOCUS | AA098755 343 bp mRNA EST 28-OCT-1996 |
| DEFINITION | T3985 MYA74 bloodstream form of serodeme WRAtat1.1 trypanosoma brucei rhodesiense cDNA 5' similar to gi 477308 pir A48583 ribosomal protein s11, mRNA sequence. |
| ACCESSION | AA098755 |
| NID | q1644707 |
| KEYWORDS | EST. |
| SOURCE | Trypanosoma brucei rhodesiense. |
| ORGANISM | Trypanosoma brucei rhodesiense Eukaryotae; mitochondrial eukaryotes; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma. 1 (bases 1 to 343) Dijkeng,A., Donelson,J.E. and Majiwa,P.A.O. Generation of expressed sequence tags as physical landmarks in the genome of Trypanosoma brucei Unpublished (1996) |
| REFERENCE | |
| AUTHORS | |
| TITLE | |
| JOURNAL | |
| COMMENT | Contact: Majiwa PAO Molecular Biology Unit International Livestock Research Institute P.O. Box 30709, Nairobi, Kenya Tel: 254-2 630743 Fax: 254-2 631499 Email: p.majiwa@cgnnet.com Seq primer: T3 primer. Location/Qualifiers 1..343 /organism="Trypanosoma brucei rhodesiense" |
| FEATURES | |
| source | |

[illegible]

Df 456 GSAGVTNVNHDWS 468
I : I : : :
Qy 2037 GGTGAAGGCACTG 2049

RESULT 13
LOCUS AF034173 2275 bp mRNA EST 22-DEC-1997
DEFINITION Homo sapiens ntcon2 contig mRNA, partial sequence, mRNA sequence.
ACCESSION AF034173
NID Q2707735
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2275)
AUTHORS Tripodis,N. and Ragousiss,J.
TITLE Generation of a transcription map in the region immediately centromeric to human MHC across the 6p21.2-6p21.3 chromosomal boundary

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2275)
AUTHORS Tripodis,N. and Ragousiss,J.
DIRECT SUBMISSION
TITLE Submitted (13-NOV-1997) Division of Medical and Molecular Genetics,
JOURNAL Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK
FEATURES Location/Qualifiers
source 1..2275 /organism='Homo sapiens'
db_xref='taxon:9606'
chromosome='6'
map='6p21.3'
clone='ntcon2 contig'
tissue_type='liver; brain'
dev_stage='fetus'
note='similar to Br140'

BASE COUNT 438 A 619 C 470 G 599 T 149 others
ORIGIN

Query Match 2.9%; Score 26; DB 11; Length 2275;
Best Local Similarity 10.0%; Pred.No.1.10e-07;
Matches 6; Conservative 37; Mismatches 17; Indels 0; Gaps 0;

Df 1483 FWRKRGRKRRMTGYMKRYRAMMAMCMMACWYWKMRGMKKCKYRKYKKTSTVV 1542
::: :::: :: :::: :::: :::: :::: :::: :::
Cc 2703 GTGGTGAGTGCAGTCACCACAGAATCATCCAAACGCCAGCGCCGTC 2644

RESULT 14
LOCUS AA848961 301 bp mRNA EST 30-APR-1998
DEFINITION EST191723 Normalized rat lung, Bento Soares Rattus sp. cDNA clone
RIUA144 3' end, mRNA sequence.
ACCESSION AA848961
NID G2936501
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 301)
AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Ketlavage,A.R. and Adams,M.D.
TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat Gene Index
JOURNAL Unpublished (1998)

Contact: Lee, NH
ATCC
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@igr.org

Search completed: Thu Dec 24 18:28:41 1998
Job time : 3855 secs.

THIS PAGE BLANK (USPTO)


```

Best Local Similarity 18.8%; Pred. No. 1.20e-10;
Matches 27; Conservative 58; Mismatches 59; Indels 0; Gaps 0;

Db    50 agnycggcgagccgacgcgaaycchqvcgymrttthhyrmrbnvyrdynrsd 109
      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Cp    2093 AGAGGCGGGGGCGCCGCCCTCGTAGTTGAGCAGCGCTGAACAGTGCCCTTACCCTC 2034

Db    110 aaawcyrrsvkydcynachdhdyvbbvbnynhhnnccccbnhvnbnh 169
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Cp    2033 GAGGTGAGAGCCTCGCCCTCTTTCTCTGCGGAGCGTTCTGCGTCCCACGAGTAGTCC 1974

Db    170 rnwayvrhdarrddvhocvchoeg 193
      : : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Cp    1973 ATGTTCAATCGCGCGCAGCCGC 1950


RESULT 4
ID     N81164 standard; DNA; 204 BP.
AC     N81164;
DT     08-NOV-1990 (first entry)
DE     Base substituted E.coli beta-galactosidase alpha-fragment.
KW     E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS     Escherichia coli.
FH     Key Location/Qualifiers
FT     misc_feature 19..69
FT     tag= a
FT     /function=multiple cloning site
FT     primer_bind 187..204
FT     tag= b
PN     EP-285123-A.
PD     05-MAY-1988.
PF     30-MAR-1988; 105163.
PR     03-APR-1987; US-034819.
PA     (SUSO) SUOMEN SOKERI OY.
PI     Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
PT     WPI; 88-279927/40.
DR     Introducing random point mutations into nucleic acids -
PT     by prepn of single stranded template, annealing a primer, elongation,
PT     misincorporation, completion of molecules and screening.
PS     Disclosure; P; English.
CC     Random point mutations were introduced into the alpha fragment of
CC     E.coli beta-galactosidase. The wild type sequence was obtained as a
CC     single stranded template and an oligonucleotide was hybridised to
CC     it to generate a popn of DNA molecules which terminate at all
CC     possible nucleotide positions within a specified region. The
CC     variable 3' ends generated in this way are used as primers for
CC     reverse transcriptase. Nucleotides are misincorporated by the
CC     transcriptase and the molecules are completed to forms that can be
CC     amplified and then expressed in a suitable host-vector system.
CC     The sequence covers all 176 diffit base substitutions, most of which
CC     occurred singularly in any given mutant.
CC     See also P80575.
SQ     Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 4.4%; Score 40; DB 1; Length 204;
Best Local Similarity 8.2%; Pred. No. 4.75e-10;
Matches 9; Conservative 59; Mismatches 42; Indels 0; Gaps 0;

Db    80 hvccgcgmrttthhyrmbnvrydnrsdaawcyrrsvkydcynachdhdyvbyb 139
      :||::||: : : : : : : : : | : : : : | : : : : | : : : :
Qy    2432 CCCTCGCTCATGAGCAGAGCTCTCCCTCCCTGAATGAGCGCAGTGCCCTTCGAGGCT 2491

Db    140 bbvynvnhnncnccbnhvhcnvbnhnrwayvrhdarrddvhocv 189
      : : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Qy    2492 TCCTAGCGTTCATGTGCCACCGCGCTGCATCATGAGGCAAGTCTTAC 2541


RESULT 5
ID     Q70465 standard; DNA; 114 BP.
AC     Q70465;
DT     05-APR-1995 (first entry)
DE     Generic DNA sequence to generate a random TSAR peptide library.
KW     TSAR; totally synthetic affinity reagent; synthetic; binding domain;

```


[illegible]

THIS PAGE BLANK (USP10)

W P S R L L (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 18:38:12 1998; MasPar time 1550.93 Seconds
Tabular output not generated. 622.512 Million cell updates/sec

Title: >US-08-951-733-13
Description: (1920-2820) from US08951733.seq (2 of 2)
Perfect Score: 901
N.A. Sequence: 1920 GTCGACTCGCTTCATCC.....AGGCCCTGGTGGCAGCGCT 2820
Comp: CAGGTCGTGAGCGAAGTAGG.....TCCGGGACCCACCGTGC

Scoring table: TABLE default
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 1665728 seqs, 535777091 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-pending
1:P9 2:U6000 3:U6001 4:U6002 5:U6003 6:U6004 7:U6005
8:U6006 9:U6007 10:U6008 11:U7 12:U80 13:U81 14:U82
15:U83 16:U84A 17:U84B 18:U85 19:U86 20:U87 21:U88
22:U89 23:U90A 24:U90B 25:U91 26:NEWP 27:NEWU6 28:NEWU8
29:NEWU9

Statistics: Mean 9.874; Variance 3.555; scale 2.777

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|----------------------------------|-----------|
| 1 | 901 | 100.0 | 2848 | 22 | US-08-951-Sequence 13, Applicati | 0.00e+00 |
| 2 | 901 | 100.0 | 3798 | 22 | US-08-951-Sequence 19, Applicati | 0.00e+00 |
| 3 | 901 | 100.0 | 4015 | 24 | US-09-052-Sequence 1, Applicati | 0.00e+00 |
| 4 | 901 | 100.0 | 4015 | 21 | US-08-854-Sequence 224, Applicat | 0.00e+00 |
| 5 | 901 | 100.0 | 4015 | 22 | US-08-912-Sequence 1, Applicati | 0.00e+00 |
| 6 | 901 | 100.0 | 4023 | 23 | US-09-026-Sequence 35, Applicati | 0.00e+00 |
| 7 | 901 | 100.0 | 7029 | 22 | US-08-911-Sequence 1, Applicati | 0.00e+00 |
| 8 | 898 | 99.7 | 4029 | 21 | US-08-854-Sequence 173, Applicat | 0.00e+00 |
| 9 | 898 | 99.7 | 4029 | 21 | US-08-851-Sequence 173, Applicat | 0.00e+00 |
| 10 | 508 | 56.4 | 2176 | 22 | US-08-912-Sequence 3, Applicati | 0.00e+00 |
| 11 | 508 | 56.4 | 3346 | 23 | US-09-026-Sequence 36, Applicati | 0.00e+00 |
| 12 | 508 | 56.4 | 3855 | 22 | US-08-911-Sequence 18, Applicati | 0.00e+00 |
| 13 | 508 | 56.4 | 3855 | 22 | US-08-912-Sequence 4, Applicati | 0.00e+00 |
| 14 | 477 | 52.9 | 2171 | 21 | US-08-851-Sequence 100, Applicat | 0.00e+00 |
| 15 | 477 | 52.9 | 2171 | 21 | US-08-854-Sequence 100, Applicat | 0.00e+00 |

| | | | | | | |
|----|-----|------|------|----|----------------------------------|-----------|
| 16 | 477 | 52.9 | 2171 | 21 | US-08-846-Sequence 100, Applicat | 0.00e+00 |
| 17 | 369 | 41.0 | 3496 | 23 | US-09-042-Sequence 52, Applicati | 3.00e-00 |
| 18 | 277 | 30.7 | 535 | 22 | US-08-911-Sequence 54, Applicati | 1.27e-231 |
| 19 | 236 | 26.2 | 550 | 22 | US-08-911-Sequence 9, Applicatio | 1.82e-169 |
| 20 | 182 | 20.2 | 182 | 22 | US-08-912-Sequence 62, Applicati | 2.84e-143 |
| 21 | 159 | 17.6 | 389 | 21 | US-08-846-Sequence 62, Applicati | 2.84e-143 |
| 22 | 159 | 17.6 | 389 | 21 | US-08-846-Sequence 62, Applicati | 2.84e-143 |
| 23 | 159 | 17.6 | 389 | 21 | US-08-851-Sequence 17, Applicati | 2.84e-143 |
| 24 | 159 | 17.6 | 389 | 22 | US-08-911-Sequence 62, Applicati | 2.84e-143 |
| 25 | 159 | 17.6 | 389 | 21 | US-08-854-Sequence 8, Applicatio | 2.84e-143 |
| 26 | 159 | 17.6 | 389 | 22 | US-08-912-Sequence 14, Applicati | 3.75e-41 |
| 27 | 66 | 7.3 | 578 | 19 | US-08-624-Sequence 14, Applicati | 3.75e-41 |
| 28 | 66 | 7.3 | 578 | 18 | US-08-569-Sequence 8, Applicatio | 5.74e-32 |
| 29 | 57 | 6.3 | 383 | 16 | US-08-446-Sequence 8, Applicatio | 5.74e-32 |
| 30 | 57 | 6.3 | 383 | 16 | US-08-446-Sequence 8, Applicatio | 5.74e-32 |
| 31 | 57 | 6.3 | 383 | 16 | US-08-446-Sequence 8, Applicatio | 5.74e-32 |
| 32 | 57 | 6.3 | 383 | 17 | US-08-446-Sequence 8, Applicatio | 5.74e-32 |
| 33 | 57 | 6.3 | 383 | 17 | US-08-451-Sequence 8, Applicatio | 5.74e-32 |
| 34 | 57 | 6.3 | 383 | 15 | US-08-311-Sequence 8, Applicatio | 5.74e-32 |
| 35 | 55 | 6.1 | 383 | 16 | US-08-446-Sequence 8, Applicatio | 5.81e-30 |
| 36 | 55 | 6.1 | 383 | 16 | US-08-446-Sequence 8, Applicatio | 5.81e-30 |
| 37 | 55 | 6.1 | 383 | 16 | US-08-446-Sequence 8, Applicatio | 5.81e-30 |
| 38 | 55 | 6.1 | 383 | 17 | US-08-451-Sequence 8, Applicatio | 5.81e-30 |
| 39 | 55 | 6.1 | 383 | 16 | US-08-446-Sequence 8, Applicatio | 5.81e-30 |
| 40 | 55 | 6.1 | 383 | 15 | US-08-311-Sequence 8, Applicatio | 5.81e-30 |
| 41 | 54 | 6.0 | 60 | 22 | US-08-911-Sequence 44, Applicati | 5.77e-29 |
| 42 | 54 | 6.0 | 60 | 22 | US-08-912-Sequence 10, Applicati | 5.77e-29 |
| 43 | 54 | 6.0 | 477 | 11 | US-07-904-Sequence 53, Applicati | 5.68e-28 |
| 44 | 53 | 5.9 | 239 | 22 | US-08-911-Sequence 14, Applicati | 5.68e-28 |
| 45 | 53 | 5.9 | 578 | 19 | US-08-624-Sequence 14, Applicati | 5.68e-28 |

ALIGNMENTS

RESULT 1
ID US-08-951-733-13 STANDARD; DNA; UNC; 2848 BP.
AC xxxxxx
DT

Sequence 13, Application US/08951733
Sequence 13, Application US/08951733
GENERAL INFORMATION:

APPLICANT: Harrington, Lea A.

APPLICANT: Robinson, Murray O.

TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Inc.

STREET: One Angen Center Drive

CITY: Thousand Oaks

STATE: CA

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/951,733

FILING DATE: 16-OCT-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/873,039

FILING DATE: 11-JUN-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/751,189

FILING DATE: 15-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Oleski, Nancy A.

REGISTRATION NUMBER: 34,688

REFERENCE/DOCKET NUMBER: A-433B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (805) 447-6504

CC TELEFAX: (805) 499-8011
CC INFORMATION FOR SEQ ID NO: 13:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2848 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
SQ SEQUENCE 2848 BP; 437 A; 978 C; 945 G; 488 T; 0 OTHER.

Query Match 100.0%; Score 901; DB 22; Length 2848;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1920 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGTGCGGCCGATTGTGAACATGGACTA 1979
Qy 1920 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGTGCGGCCGATTGTGAACATGGACTA 1979

Db 1980 CBTCTGGGAGCCAGACGTTCCGACAGAGAAAGAGGGCCGAGCGTCTCACCTCGAGGTT 2039
Qy 1980 CBTCTGGGAGCCAGACGTTCCGACAGAGAAAGAGGGCCGAGCGTCTCACCTCGAGGTT 2039

Db 2040 GAAGGCACTGTTACGCGTCTCAACTACGAGGGGGCGGCCCGCTCTCTGGGGCC 2099
Qy 2040 GAAGGCACTGTTACGCGTCTCAACTACGAGGGGGCGGCCCGCTCTCTGGGGCC 2099

Db 2100 CTCTGTGTCGGCCCTGGACGATATCCACAGGGCCTGGCGACCTTCGTGCTGTGGG 2159
Qy 2100 CTCTGTGTCGGCCCTGGACGATATCCACAGGGCCTGGCGACCTTCGTGCTGTGGG 2159

Db 2160 GCGCCAGGACCGCGCCCTGAGCTGTACTTTGTCAGGTGGATGTGAGGGGGCGGTAGCA 2219
Qy 2160 GCGCCAGGACCGCGCCCTGAGCTGTACTTTGTCAGGTGGATGTGAGGGGGCGGTAGCA 2219

Db 2220 CACCATCCCCAGGACAGCTCACGGAGTTCATCCGACGATCATCAACCCAGAACAC 2279
Qy 2220 CACCATCCCCAGGACAGCTCACGGAGTTCATCCGACGATCATCAACCCAGAACAC 2279

Db 2280 GTACTGCGTGCCTGCTGATGCGGTGCTCCAGAGGCGGCCATGGCAGCTCCGCAAGGC 2339
Qy 2280 GTACTGCGTGCCTGCTGATGCGGTGCTCCAGAGGCGGCCATGGCAGCTCCGCAAGGC 2339

Db 2340 CTTCAAGAGCCACGCTCTACCTTCACAGACCTCCAGCGTACATGCCACAGTTGCTGGC 2399
Qy 2340 CTTCAAGAGCCACGCTCTACCTTCACAGACCTCCAGCGTACATGCCACAGTTGCTGGC 2399

Db 2400 TCACCTGCAGGACAGACCGCCCTGAGGGATGCCGTGCTCATCGACGAGCTCTCCCT 2459
Qy 2400 TCACCTGCAGGACAGACCGCCCTGAGGGATGCCGTGCTCATCGACGAGCTCTCCCT 2459

Db 2460 GAATGAGGCCAGCTGCGCTCTTCGACGCTTCCTACGCTTCATCGACGAGCTCCACCGCT 2519
Qy 2460 GAATGAGGCCAGCTGCGCTCTTCGACGCTTCCTACGCTTCATCGACGAGCTCCACCGCT 2519

Db 2520 GCGCATCAGGGCAAGTCTTAGTCCAGTCCAGGGATCCCGAGGGCTCCATCCTCTC 2579
Qy 2520 GCGCATCAGGGCAAGTCTTAGTCCAGTCCAGGGATCCCGAGGGCTCCATCCTCTC 2579

Db 2580 CACGCTGCTCTCAGCCTGTGCTACGGGACATGGAGAACAGCTGTTGGGGGATTCG 2639
Qy 2580 CACGCTGCTCTCAGCCTGTGCTACGGGACATGGAGAACAGCTGTTGGGGGATTCG 2639

Db 2640 GCGGACCGGCTGCTCTCCTGCTGATGATTTCTGTTGGTGACACCTCACCTCAC 2699
Qy 2640 GCGGACCGGCTGCTCTCCTGCTGATGATTTCTGTTGGTGACACCTCACCTCAC 2699

Db 2700 CCACCGGAAACCTTCCTCAGACCCCTGGTCGAGGTGCTCCCTGAGTATGGCTGGGT 2759
Qy 2700 CCACCGGAAACCTTCCTCAGACCCCTGGTCGAGGTGCTCCCTGAGTATGGCTGGGT 2759

Db 2760 GAACTTGGGGAAGACAGTGGTGAACCTCCCTGTAGAAGACAGGGCCCTGGGTGGCACGGC 2819
Qy 2760 GAACTTGGGGAAGACAGTGGTGAACCTCCCTGTAGAAGACAGGGCCCTGGGTGGCACGGC 2819

Db 2820 T 2820
Qy 2820 T 2820

RESULT 2
ID US-08-951-733-19 STANDARD; DNA; UNC; 3798 BP.
AC xxxxxx
DT
DE Sequence 19, Application US/08951733
CC Sequence 19, Application US/08951733
CC GENERAL INFORMATION:
CC APPLICANT: Harrington, Lea A.
CC APPLICANT: Robinson, Murray O.
CC TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS
CC NUMBER OF SEQUENCES: 44
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Amgen Inc.
CC STREET: One Amgen Center Drive
CC CITY: Thousand Oaks
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 91320-1789
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/951,733
CC FILING DATE: 16-OCT-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/873,039
CC FILING DATE: 11-JUN-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/751,189
CC FILING DATE: 15-NOV-1996
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Oleski, Nancy A.
CC REGISTRATION NUMBER: 34,688
CC REFERENCE/DOCKET NUMBER: A-433B
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (805) 447-6504
CC TELEFAX: (805) 499-8011
CC INFORMATION FOR SEQ ID NO: 19:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 3798 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
SQ SEQUENCE 3798 BP; 613 A; 1310 C; 1213 G; 562 T; 0 OTHER.

Query Match 100.0%; Score 901; DB 22; Length 3798;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1921 GTCAGAGCTCGCTTCATCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTA 1980
Qy 1921 GTCAGAGCTCGCTTCATCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTA 1979

Db 1981 CGTCGTGGGAGCCAGACGTTCCGACAGAGAAAGAGGGCCGAGCGTCTCACCTCGAGGTT 2040
Qy 1981 CGTCGTGGGAGCCAGACGTTCCGACAGAGAAAGAGGGCCGAGCGTCTCACCTCGAGGTT 2039

Db 2041 GAAGGCACTGTTACGCGTCTCAACTACGAGGGGGCGGCCCGCTCTCTGGGGCC 2100
Qy 2041 GAAGGCACTGTTACGCGTCTCAACTACGAGGGGGCGGCCCGCTCTCTGGGGCC 2099

Db 2101 CTCTGTGCTGGGCTTGACGATATCCACAGGGCTGCGGACCTTCGTGCTGCTGTGGC 2160
Qy 2101 CTCTGTGCTGGGCTTGACGATATCCACAGGGCTGCGGACCTTCGTGCTGCTGTGGC 2160

QY 2100 CTCTGTGCTGGCGCTGACGATATCCACAGGGCTGCGCACCTTCGTGCTGCTGTGG 2159
Db 2161 GGCCCGAGACCCCGCCTGAGCTGACTTTGTCAAGGTGGATGTACGGGCGCGTACGA 2220
QY 2160 GGCCCGAGACCCCGCCTGAGCTGACTTTGTCAAGGTGGATGTACGGGCGCGTACGA 2219
Db 2221 CACCATCCCCCAGCAGAGGCTCAGGAGGTCTATGCCAGCATCATCAAAACCCAGAACAC 2280
QY 2220 CACCATCCCCCAGCAGAGGCTCAGGAGGTCTATGCCAGCATCATCAAAACCCAGAACAC 2279
Db 2281 GTACTGCGTGGCTGCTATGCCGTGTGTCCAGAGGGCGCCCATGGGCACCTCCGCAAGGC 2340
QY 2280 GTACTGCGTGGCTGCTATGCCGTGTGTCCAGAGGGCGCCCATGGGCACCTCCGCAAGGC 2339
Db 2341 CTTCAAGAGCCAGCTCTTACCTTTGACAGACCTCCAGCCGCTATCGCAGAGTTCGTGGC 2400
QY 2340 CTTCAAGAGCCAGCTCTTACCTTTGACAGACCTCCAGCCGCTATCGCAGAGTTCGTGGC 2399
Db 2401 TCACCTGCAGGAGACAGCCGCTGAGGGATGCCGTGTCATCGACGAGAGCTCCTCCCT 2460
QY 2400 TCACCTGCAGGAGACAGCCGCTGAGGGATGCCGTGTCATCGACGAGAGCTCCTCCCT 2459
Db 2461 GAATGAGGCCAGCAGTGGCCTCTTCGACGCTTCTTACGCTTCATGTGCCACACGCGCT 2520
QY 2460 GAATGAGGCCAGCAGTGGCCTCTTCGACGCTTCTTACGCTTCATGTGCCACACGCGCT 2519
Db 2521 GCGCATCAGGGGAAGTCTTACGTTCAGTCCAGTCCAGGGGATCCCGAGGGTCCATCTCTC 2580
QY 2520 GCGCATCAGGGGAAGTCTTACGTTCAGTCCAGTCCAGGGGATCCCGAGGGTCCATCTCTC 2579
Db 2581 CAGCTGCTCTGAGCTGTGCTACGCGGACATGGAGAAAGCTGTTGCGGGGATTCG 2640
QY 2580 CAGCTGCTCTGAGCTGTGCTACGCGGACATGGAGAAAGCTGTTGCGGGGATTCG 2639
Db 2641 GCGGGAGCGGCTGCTCTGCTGTTTGGTGGATGATTTCTTGTGTGACACTCACCTCAC 2700
QY 2640 GCGGGAGCGGCTGCTCTGCTGTTTGGTGGATGATTTCTTGTGTGACACTCACCTCAC 2699
Db 2701 CCACGGGAAAACCTTCTCAGGACCTGCTCGAGGTGTCCTCGAGTATGCTGCTGGT 2760
QY 2700 CCACGGGAAAACCTTCTCAGGACCTGCTCGAGGTGTCCTCGAGTATGCTGCTGGT 2759
Db 2761 GAATTCGGAAGACAGTGTGACTTCCCTGTAGAGAGAGGCGCTGGGTGGCAGGC 2820
QY 2760 GAATTCGGAAGACAGTGTGACTTCCCTGTAGAGAGAGGCGCTGGGTGGCAGGC 2819
Db 2821 T 2821
QY 2820 T 2820

RESULT 3
ID US-09-052-919-1 STANDARD; DNA; UNC; 4015 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/09052919
CC Sequence 1, Application US/09052919
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Antisense Compositions for Detecting and
CC TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase
CC NUMBER OF SEQUENCES: 72
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California

CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/052,919
CC FILING DATE: 31-MAR-1998
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/974,549
CC FILING DATE: 19-NOV-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/974,584
CC FILING DATE: 19-NOV-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Parent, Annette S.
CC REGISTRATION NUMBER: 42,058
CC REFERENCE/DOCKET NUMBER: 015389-00360005
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4015 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 56..3454
CC OTHER INFORMATION: /product= "human telomerase reverse
CC OTHER INFORMATION: transcriptase (hMRT)"
CC
SQ SEQUENCE 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 OTHER.

Query Match 100.0%; Score 901; DB 24; Length 4015;
Best Local Similarity 100.0%; Pred.No. 0.00e+00;
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1909 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGCTGCGGCCGACTTGTGAACATGGACTA 1968

QY 1920 GTCCAGACTCGCTTCATCCCAAGCTGACGGCTGCGGCGATTGTGAACATGGACTA 1979
Db 1969 CGTCGTGGAGCCAGAACGTTCCGACAGAAAAGAGGCGCGGTCTCACCTCGAGGGT 2028
QY 1980 CGTCGTGGAGCCAGAACGTTCCGACAGAAAAGAGGCGCGGTCTCACCTCGAGGGT 2039
Db 2029 GAAGGCACTGTTGAGCGTCTCACTAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2088
QY 2040 GAAGGCACTGTTGAGCGTCTCACTAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2099
Db 2089 CTCTGTGCTGGGCGCTGACGATATCCAGGGCGCTGCGGCGGCGGCGGCGGCGGCGG 2148
QY 2100 CTCTGTGCTGGGCGCTGACGATATCCAGGGCGCTGCGGCGGCGGCGGCGGCGGCGG 2159
Db 2149 GGCCAGGACCCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGCGCGGTACGA 2208
QY 2160 GGCCAGGACCCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGCGCGGTACGA 2219
Db 2209 CACCATCCCGACAGACAGGCTCAGGAGGTGATCGGAGCATATCAAAACCCCAAGAAC 2268
QY 2220 CACCATCCCGACAGACAGGCTCAGGAGGTGATCGGAGCATATCAAAACCCCAAGAAC 2279
Db 2269 GTACTGCTGCTGGTATGCGTGTGCTCCAGAGGCGCGGCGGCGGCGGCGGCGGCGG 2328
QY 2280 GTACTGCTGCTGGTATGCGTGTGCTCCAGAGGCGCGGCGGCGGCGGCGGCGGCGG 2339
Db 2329 CTTCAAGAGCCACGCTCTACTTTGACAGACCTCCAGCGGCTACATGCGAGAGTTCTG 2388
QY 2340 CTTCAAGAGCCACGCTCTACTTTGACAGACCTCCAGCGGCTACATGCGAGAGTTCTG 2399
Db 2389 TCACCTCGAGGAGACAGCGGCTGAGGATGCGGCTGATCGAGAGAGCTCCTCCCT 2448
QY 2400 TCACCTCGAGGAGACAGCGGCTGAGGATGCGGCTGATCGAGAGAGCTCCTCCCT 2459
Db 2449 GAATGAGCCAGCAGTGGCTCTTCGACGCTTCTTACGCTTCTTACGCTTCTTACG 2508
QY 2460 GAATGAGCCAGCAGTGGCTCTTCGACGCTTCTTACGCTTCTTACGCTTCTTACG 2519
Db 2509 GCGCATCAGGCGCAAGCTCTACGCTTCCAGTCCAGGCGGATCCCGAGGCGCTCCATCTC 2568
QY 2520 GCGCATCAGGCGCAAGCTCTACGCTTCCAGTCCAGGCGGATCCCGAGGCGCTCCATCTC 2579
Db 2569 CACGCTGCTGAGCGCTGCTAGGCGGATGCGGCGGATGCGGCGGATGCGGCGGATGCG 2628
QY 2580 CACGCTGCTGAGCGCTGCTAGGCGGATGCGGCGGATGCGGCGGATGCGGCGGATGCG 2639
Db 2629 GCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2688
QY 2640 GCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2699
Db 2689 CCACGCGAAACCTTCTCAGGACCTTGGTCCGAGGTGCTCCCTGAGTATGGCTGCGTGGT 2748
QY 2700 CCACGCGAAACCTTCTCAGGACCTTGGTCCGAGGTGCTCCCTGAGTATGGCTGCGTGGT 2759
Db 2749 GAATTCGCGAAGACATGGTGAATCTCCCTGTAGAGAGAGGCGGCTGGGTGCGACGCG 2808
QY 2760 GAATTCGCGAAGACATGGTGAATCTCCCTGTAGAGAGAGGCGGCTGGGTGCGACGCG 2819
Db 2809 T 2809
QY 2820 T 2820

RESULT 4
ID US-08-854-050-224 STANDARD; DNA; UNC; 4015 BP.
AC xxxxxx

DT Sequence 224, Application US/08854050
CC Sequence 224, Application US/08854050
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Novel Telomerase
CC NUMBER OF SEQUENCES: 225
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-0029300S
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 224:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4015 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 56..3454
CC OTHER INFORMATION: /product= "hTERT"
CC OTHER INFORMATION: /note= "human telomerase reverse
CC OTHER INFORMATION: transcriptase (hTERT) catalytic protein
CC OTHER INFORMATION: component"
CC SEQUENCE 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 OTHER.

Query Match 100.0%; Score 901; DB 21; Length 4015;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1909 GTCCAGACTCGGCTTCATCCCAAGCTGACGGGCTGCGGCGGATGTGAACATGGACTA 1968
QY 1920 GTCCAGACTCGGCTTCATCCCAAGCTGACGGGCTGCGGCGGATGTGAACATGGACTA 1979
Db 1969 CGTCGTGGAGCCAGAACGTTCCGACAGAAAAGAGGCGGCGGCGGCGGCTCACCCTCGAGGGT 2028
QY 1980 CGTCGTGGAGCCAGAACGTTCCGACAGAAAAGAGGCGGCGGCGGCTCACCCTCGAGGGT 2039

Db 2029 GAAGGCACTGTTACAGCGTCTCACTACGAGCGGGCGGGCCCGCCGCTCTCTGGGGCG 2088
Qy 2040 GAAGGCACTGTTACAGCGTCTCACTACGAGCGGGCGGGCCCGCCGCTCTCTGGGGCG 2099
Db 2089 CTCTGCTGCTGGGCGCTGAGAGATATCCACAGGGCGCTGGCGACCTTCGTCGTCGTCGTCG 2148
Qy 2100 CTCTGCTGCTGGGCGCTGAGAGATATCCACAGGGCGCTGGCGACCTTCGTCGTCGTCGTCG 2159
Db 2149 GGGCCAGACCGCGCGCTGAGCTGCTACTTTGTCAGGTGGATGTGACGGGCGGTACGA 2208
Qy 2160 GGGCCAGACCGCGCGCTGAGCTGCTACTTTGTCAGGTGGATGTGACGGGCGGTACGA 2219
Db 2209 CACCATCCCCAGGACAGGCTCACGGAGGTATCCCGACGATCATCAAAACCCAGAACAC 2268
Qy 2220 CACCATCCCCAGGACAGGCTCACGGAGGTATCCCGACGATCATCAAAACCCAGAACAC 2279
Db 2269 GTACTGCGTGGCTGCGTATGCGGTGGTCCAGAAAGCGCCGATGGCGACGTCCCGAAGGC 2328
Qy 2280 GTACTGCGTGGCTGCGTATGCGGTGGTCCAGAAAGCGCCGATGGCGACGTCCCGAAGGC 2339
Db 2329 CTTCAAGAGCCACGCTCTTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTGCTGGC 2388
Qy 2340 CTTCAAGAGCCACGCTCTTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTGCTGGC 2399
Db 2389 TCACCTGACGAGACACCGCGCTGAGGATGCCGTGCTCATCGAGCAGAGCTCTCCCT 2448
Qy 2400 TCACCTGACGAGACACCGCGCTGAGGATGCCGTGCTCATCGAGCAGAGCTCTCCCT 2459
Db 2449 GAATGAGCCACGAGTGGCTCTTCGAGCTCTTCCTACGCTCTTCCTACGCTTCATGTCGACACGCGGT 2508
Qy 2460 GAATGAGCCACGAGTGGCTCTTCGAGCTCTTCCTACGCTCTTCCTACGCTTCATGTCGACACGCGGT 2519
Db 2509 GCGCATCAGGGCAAGTCTTACCTGACAGACCTCCAGCCGTACATGCGACAGTTGCTGGC 2568
Qy 2520 GCGCATCAGGGCAAGTCTTACCTGACAGACCTCCAGCCGTACATGCGACAGTTGCTGGC 2579
Db 2569 CACGCTGCTTCGACCGCTGTGTCAGGGGACATGAGAACAGCTGTTTGGGGGATTCG 2628
Qy 2580 CACGCTGCTTCGACCGCTGTGTCAGGGGACATGAGAACAGCTGTTTGGGGGATTCG 2639
Db 2629 GGGGACCGGCTGCTCTGCGCTTTCGCTGATGATTTCTTGTGTGACACCTCACCTCAC 2688
Qy 2640 GGGGACCGGCTGCTCTGCGCTTTCGCTGATGATTTCTTGTGTGACACCTCACCTCAC 2699
Db 2699 CCACCGGCAAAACCTTCTCAGGACCTGCTCGAGGTGCTCCGTAGTATGCTGCGTGGT 2748
Qy 2700 CCACCGGCAAAACCTTCTCAGGACCTGCTCGAGGTGCTCCGTAGTATGCTGCGTGGT 2759
Db 2749 GAACCTGCGGAGACAGTGGTGAATTCCTCTGTAAGACGAGCCCTGGGTGGCACGGC 2808
Qy 2760 GAACCTGCGGAGACAGTGGTGAATTCCTCTGTAAGACGAGCCCTGGGTGGCACGGC 2819
Db 2809 T 2809
Qy 2820 T 2820

RESULT 5
ID US-08-912-951-1 STANDARD; DNA; UNC; 4015 BP.
AC xxxxx

DE Sequence 1, Application US/08912951
CC Sequence 1, Application US/08912951
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND

CC TITLE OF INVENTION: THERAPEUTIC METHODS
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912.951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002600US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4015 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 56..3454
CC OTHER INFORMATION: /product= "htrr"
CC OTHER INFORMATION: /note= "human telomerase reverse
CC OTHER INFORMATION: transcriptase (htrr) catalytic protein
CC OTHER INFORMATION: component"
CC SEQUENCE 4015 BP: 663 A; 1363 C; 1275 G; 714 T; 0 OTHER.

Query Match 100.0%; Score 901; DB 22; Length 4015;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1909 GTCCAGACTCGCTTCATCCCAAGCCCTGACGGGCTGGCGCCGATGTTGAACATGGACTA 1968
Qy 1920 GTCCAGACTCGCTTCATCCCAAGCCCTGACGGGCTGGCGCCGATGTTGAACATGGACTA 1979
Db 1969 CGTCGTGGGAGCCAGACGTTCCCGAGAGAAAGGGCCGAGCGGTCTCACCTCGAGGGT 2028
Qy 1980 CGTCGTGGGAGCCAGACGTTCCCGAGAGAAAGGGCCGAGCGGTCTCACCTCGAGGGT 2039

| | | | |
|--------|---------------------------|---|------|
| Db | 2444 | GAATGAGGCCAGAGTGGCCTCTTCGACGCTTCTACGCTTCTATGTGCGCCACACGCGGT | 2503 |
| Qy | 2460 | GAATGAGGCCAGAGTGGCCTCTTCGACGCTTCTACGCTTCTATGTGCGCCACACGCGGT | 2519 |
| Db | 2504 | GCGCATCAGGGGCAAGTCCCTACGTCAGTCGACGCGGATCCCGACGAGGCTCCATCCTCTC | 2563 |
| Qy | 2520 | GCGCATCAGGGGCAAGTCCCTACGTCAGTCGACGCGGATCCCGACGAGGCTCCATCCTCTC | 2579 |
| Db | 2564 | CACGCTGCTCTGCAGCCTGTGCTACGGGACATGGAGAACAGCTGTTTGGCGGGATTCTG | 2623 |
| Qy | 2580 | CACGCTGCTCTGCAGCCTGTGCTACGGGACATGGAGAACAGCTGTTTGGCGGGATTCTG | 2639 |
| Db | 2624 | GCGGGACGGGCTGCTCCTCGCTTGGTGGATGATTTCTGTGTGTGACACCTCACCTCAC | 2683 |
| Qy | 2640 | GCGGGACGGGCTGCTCCTCGCTTGGTGGATGATTTCTGTGTGTGACACCTCACCTCAC | 2699 |
| Db | 2684 | CCAGCGGAAACCTTCTCTCAGAACCTTGGTCGAGAGTGTCCCTGAGTATGGCTCGTGCT | 2743 |
| Qy | 2700 | CCAGCGGAAACCTTCTCTCAGAACCTTGGTCGAGAGTGTCCCTGAGTATGGCTCGTGCT | 2759 |
| Db | 2744 | GAACTTGGGAAGACAGTGGTGAACCTTCCTCTGAGAACGAGGCGCTTGGGTGGCAGCGC | 2803 |
| Qy | 2760 | GAACTTGGGAAGACAGTGGTGAACCTTCCTCTGAGAACGAGGCGCTTGGGTGGCAGCGC | 2819 |
| Db | 2804 | T 2804 | |
| Qy | 2820 | T 2820 | |
| RESULT | 9 | | |
| ID | US-08-851-843-173 | STANDARD; DNA; UNC; 4029 BP. | |
| AC | xxxxxx | | |
| DT | | | |
| DE | Sequence 173, | Application US/08851843 | |
| CC | Sequence 173, | Application US/08851843 | |
| CC | GENERAL INFORMATION: | | |
| CC | APPLICANT: | Cech, Thomas R. | |
| CC | APPLICANT: | Lingner, Joachim | |
| CC | APPLICANT: | Nakamura, Toru | |
| CC | APPLICANT: | Chapman, Karen B. | |
| CC | APPLICANT: | Morin, Gregg B. | |
| CC | APPLICANT: | Harley, Calvin | |
| CC | APPLICANT: | Andrews, William H. | |
| CC | TITLE OF INVENTION: | Novel Telomerase | |
| CC | NUMBER OF SEQUENCES: | 223 | |
| CC | CORRESPONDENCE ADDRESS: | | |
| CC | ADDRESS: | Townsend and Townsend and Crew LLP | |
| CC | STREET: | Two Embarcadero Center, 8th Floor | |
| CC | CITY: | San Francisco | |
| CC | STATE: | California | |
| CC | COUNTRY: | United States of America | |
| CC | ZIP: | 94111 | |
| CC | COMPUTER READABLE FORM: | | |
| CC | MEDIUM TYPE: | Floppy disk | |
| CC | COMPUTER: | IBM PC compatible | |
| CC | OPERATING SYSTEM: | PC-DOS/MS-DOS | |
| CC | SOFTWARE: | PatentIn Release #1.0, Version #1.30 | |
| CC | CURRENT APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: | US/08/851,843 | |
| CC | FILING DATE: | 06-MAY-1997 | |
| CC | CLASSIFICATION: | 536 | |
| CC | PRIOR APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: | US 08/846,017 | |
| CC | FILING DATE: | 25-APR-1997 | |
| CC | CLASSIFICATION: | 536 | |
| CC | PRIOR APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: | US 08/844,419 | |
| CC | FILING DATE: | 18-APR-1997 | |
| CC | CLASSIFICATION: | 536 | |
| CC | PRIOR APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: | US 08/724,643 | |
| CC | FILING DATE: | 01-OCT-1996 | |

| | | |
|----|--------------------------------|--|
| CC | CLASSIFICATION: | 536 |
| CC | ATTORNEY/AGENT INFORMATION: | |
| CC | NAME: | Apple, Randolph T. |
| CC | REGISTRATION NUMBER: | 36,429 |
| CC | REFERENCE/DOCKET NUMBER: | 015389-002930US |
| CC | TELECOMMUNICATION INFORMATION: | |
| CC | TELEPHONE: | (415) 576-0200 |
| CC | TELEFAX: | (415) 576-0300 |
| CC | INFORMATION FOR SEQ ID NO: | 173: |
| CC | SEQUENCE CHARACTERISTICS: | |
| CC | LENGTH: | 4029 base pairs |
| CC | TYPE: | nucleic acid |
| CC | STRANDEDNESS: | single |
| CC | TOPOLOGY: | linear |
| CC | MOLECULE TYPE: | CDNA |
| CC | FEATURE: | |
| CC | NAME/KEY: | |
| CC | LOCATION: | 1..4029 |
| CC | OTHER INFORMATION: | [note= "preliminary sequence for |
| CC | OTHER INFORMATION: | human TRT CDNA insert of |
| CC | OTHER INFORMATION: | plasmid pGRN121" |
| CC | SEQUENCE | 4029 BP; 687 A; 1342 C; 1255 G; 720 T; 25 OTHER. |
| SQ | Query Match | 99.7%; Score 898; DB 21; Length 4029; |
| | Best Local Similarity | 99.8%; Pred. No. 0.00e+00; |
| | Matches | 899; Conservative 0; Mismatches 2; Indels 0; Gaps 0; |
| Df | 1904 | GTCAGACTCGCTTCATCCCAGCCTGACGGGCTCGGCCGATTGTGAACATGGACTA 1963 |
| Yq | 1920 | GTCAGACTCGCTTCATCCCAGCCTGACGGGCTCGGCCGATTGTGAACATGGACTA 1979 |
| Df | 1964 | CGTCGTGGGACCAGAAGCTTCCGACAGAAAGAGGGCGGAGCGTCTCACCTCGAGGGT 2023 |
| Yq | 1980 | CGTCGTGGGACCAGAAGCTTCCGACAGAAAGAGGGCGGAGCGTCTCACCTCGAGGGT 2039 |
| Df | 2024 | GAAGGCATGTTACAGCTGTCTCAACTACGAGGCGCGGCGCCCGCTCTCTGGGCGC 2083 |
| Yq | 2040 | GAAGGCATGTTACAGCTGTCTCAACTACGAGGCGCGGCGCCCGCTCTCTGGGCGC 2099 |
| Df | 2084 | CTCTGTCTGGGCTGGACCATATCCACAGGGCTGCGGCACACTTCTGCTGGTGTCG 2143 |
| Yq | 2100 | CTCTGTCTGGGCTGGACCATATCCACAGGGCTGCGGCACACTTCTGCTGGTGTCG 2159 |
| Df | 2144 | GGCCCCAGGACCGCGGCTGAGCTGTACTTTGTCAAAGTGGATGTGACGGGCGCGTACGA 2203 |
| Yq | 2160 | GGCCCCAGGACCGCGGCTGAGCTGTACTTTGTCAAAGTGGATGTGACGGGCGCGTACGA 2219 |
| Df | 2204 | CACCATCCCCAGGACAGGCTCAGGAGGTATCGCCAGCATCATCAAAACCCAGAACAC 2263 |
| Yq | 2220 | CACCATCCCCAGGACAGGCTCAGGAGGTATCGCCAGCATCATCAAAACCCAGAACAC 2279 |
| Df | 2264 | GTACTGGTGGTGGTATGCCGTGGTCCAGAGGGCGCCCATGGGACGCTCGCCGACGC 2323 |
| Yq | 2280 | GTACTGGTGGTGGTATGCCGTGGTCCAGAGGGCGCCCATGGGACGCTCGCCGACGC 2339 |
| Df | 2324 | CTTCAAGAGCCACGCTCTACTTTGACAGACCTCCAGCCGTACATGGACAGTTCGTGGC 2383 |
| Yq | 2340 | CTTCAAGAGCCACGCTCTACTTTGACAGACCTCCAGCCGTACATGGACAGTTCGTGGC 2399 |
| Df | 2384 | TCACCTCGAGSANAACCGCGCTGAGGGATGCGCGTCTCATCGAGCAGAGCTCTCCCT 2443 |
| Yq | 2400 | TCACCTCGAGSAGACACCGCGCTGAGGGATGCGCGTCTCATCGAGCAGAGCTCTCCCT 2459 |
| Df | 2444 | GAATGAGCCAGCAGTGGCCTTTCGACGCTTCTCCTACGCTTCAATGTGCCACACGCGGT 2503 |
| Yq | 2460 | GAATGAGCCAGCAGTGGCCTTTCGACGCTTCTCCTACGCTTCAATGTGCCACACGCGGT 2519 |
| Df | 2504 | GCGCATCAGGGCAAGTCTCTACGTCCAGTCCAGGGGATCCCGCAGGGCTCCATCTCTC 2563 |
| Yq | 2520 | GCGCATCAGGGCAAGTCTCTACGTCCAGTCCAGGGGATCCCGCAGGGCTCCATCTCTC 2579 |
| Df | 2564 | CACGCTGCTCTGCAGCCTGTGCTACGGCGACATGGAGAACAAAGCTGTTTGGGGATTGG 2623 |

QY 2580 CACGCTGCTGACGCTGTGCTAGCGGACATGAGACAAAGCTGTTGGGGGATTCG 2639

Db 2624 GCGGACGGGCTGCTCCGCTTTGGTGATGATTTCTTTGGTGACACCTCACCTCAC 2683

QY 2640 GCGGACGGGCTGCTCCGCTTTGGTGATGATTTCTTTGGTGACACCTCACCTCAC 2699

Db 2684 CACGCGAAACCTTCTCAGACCGCTGGTCCGAGGTGCTCCTGAGTATGCTGGTGGT 2743

QY 2700 CCACGCGAAACCTTCTCAGACCGCTGGTCCGAGGTGCTCCTGAGTATGCTGGTGGT 2759

Db 2744 GAACCTGCGGAAGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGTGCGACGGC 2803

QY 2760 GAACCTGCGGAAGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGTGCGACGGC 2819

Db 2804 T 2804

QY 2820 T 2820

RESULT 10

AC US-08-912-951-3 STANDARD; DNA; UNC; 2176 BP.

DT xxxxxx

DE Sequence 3, Application US/08912951

CC Sequence 3, Application US/08912951

CC GENERAL INFORMATION:

CC APPLICANT: Cech, Thomas R.

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin

CC APPLICANT: Andrews, William H.

CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND

CC TITLE OF INVENTION: THERAPEUTIC METHODS

CC NUMBER OF SEQUENCES: 335

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Townsend and Townsend and Crew LLP

CC STREET: Two Embarcadero Center, 8th Floor

CC CITY: San Francisco

CC STATE: California

CC COUNTRY: United States of America

CC ZIP: 94111

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent in Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/912,951

CC FILING DATE: 14-AUG-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/854,050

CC FILING DATE: 09-MAY-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/851,843

CC FILING DATE: 06-MAY-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/846,017

CC FILING DATE: 25-APR-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/844,419

CC FILING DATE: 18-APR-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/724,643

CC FILING DATE: 01-OCT-1996

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Apple, Randolph T.

CC REGISTRATION NUMBER: 36,429

CC REFERENCE/DOCKET NUMBER: 015389-002600US

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (415) 576-0200

CC TELEFAX: (415) 576-0300

CC INFORMATION FOR SEQ ID NO: 3:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 2176 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: cDNA

CC FEATURE:

CC NAME/KEY: -

CC LOCATION: 1..2176

CC OTHER INFORMATION: /note= "clone 712562"

CC SEQUENCE 2176 BP; 432 A; 678 C; 642 G; 422 T; 2 OTHER.

Query Match 56.4%; Score 508; DB 22; Length 2176;

Best Local Similarity 79.8%; Pred. No. 0.00e+00;

Matches 719; Conservative 0; Mismatches 0; Indels 182; Gaps 1;

Db 232 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGCTCGGCGCCGATTGTGAACATGGACTA 291

QY 1920 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGCTCGGCGCCGATTGTGAACATGGACTA 1979

Db 292 CGTCGTGGAGCCAGAACGTTCCGACAGAAAGAGGCCGAGCGTCTCACTCCGAGGGT 351

QY 1980 CGTCGTGGAGCCAGAACGTTCCGACAGAAAGAGGCCGAGCGTCTCACTCCGAGGGT 2039

Db 352 GAAGGCACCTGTTCAAGCGTGTCTCAACTACGAGCGGGCGCGCCGCCCTCTCTGGCGC 411

QY 2040 GAAGGCACCTGTTCAAGCGTGTCTCAACTACGAGCGGGCGCGCCGCCCTCTCTGGCGC 2099

Db 412 CTCTGTCTGGCGCTGGACGATATCCACAGGGCTGGCGCACCTTCGTGTGTGGTGTGG 471

QY 2100 CTCTGTCTGGCGCTGGACGATATCCACAGGGCTGGCGCACCTTCGTGTGTGGTGTGG 2159

Db 472 GGCCAGGACCGCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTACGGCGCGGTACGA 531

QY 2160 GGCCAGGACCGCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTACGGCGCGGTACGA 2219

Db 532 CACATCCCCAGACAGGCTCAGGAGTCTATCGCAGCATCATCAAAACCCAGAAAC 591

QY 2220 CACATCCCCAGACAGGCTCAGGAGTCTATCGCAGCATCATCAAAACCCAGAAAC 2279

Db 592 GTACTGCGTGGTGGTATGCGGTGTCAGAGAGGCCGCCATGGGCACGTCCGCAAGGC 651

QY 2280 GTACTGCGTGGTGGTATGCGGTGTCAGAGAGGCCGCCATGGGCACGTCCGCAAGGC 2339

Db 652 CTTCAAGAGCCACGTC----- 667

QY 2340 CTTCAAGAGCCACGTCCTCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGC 2399

Db 667 ----- 667

QY 2400 TCACCTGCAGGAGACACGCGCGCTGAGGGATGCGGTGCTCATCGAGCAGAGCTCCTCCCT 2459

Db 667 ----- 667

QY 2460 GAATGAGCCAGCAGTGGCCTCTTCGACGCTTCCTTACGCTTCATGTGCCACACGCCGT 2519

Db 668 -----CTACGTCCAGTGCAGGGGATCCCGCAGGGCTCCATCCCTCTC 709

QY 2520 GCGCATCAGGGGCAAGTCTACGTCCAGTGCAGGGGATCCCGCAGGGCTCCATCCCTCTC 2579

Db 710 CACGCTCTGTGAGCCTGTGTCTACGGCGACATGGAGAACAAAGCTGTTGCGGGGATTCG 769

QY 2580 CACGCTCTGTGAGCCTGTGTCTACGGCGACATGGAGAACAAAGCTGTTGCGGGGATTCG 2639

Db 770 GCGGACGGGCTGCTCCTCGTGGTGGATGATTTCTTCTGTGTGACACTCACCTCAC 829

QY 2640 GCGGACGGGCTGCTCTCGCTTTGGTGATGATTTCTTTGGTGACACCTCACCTCAC 2699
Db 830 CCACGCGAAACCTTCTCAGACACCTGGTCCGAGGTGTCCCTGAGTATGCTGGTGGT 889
QY 2700 CCACGCGAAACCTTCTCAGACACCTGGTCCGAGGTGTCCCTGAGTATGCTGGTGGT 2759
Db 890 GAACCTGGGGAAGACAGTGGTGAACCTTCCCTGTAGAACACGAGGCCCTGGGTGACAGGC 949
QY 2760 GAACCTGGGGAAGACAGTGGTGAACCTTCCCTGTAGAACACGAGGCCCTGGGTGACAGGC 2819
Db 950 T 950
QY 2820 T 2820

RESULT 11

ID US-09-026-981-36 STANDARD; DNA; UNC; 3346 BP.
AC xxxxxx
DT
DE Sequence 36, Application US/09026981
CC Sequence 36, Application US/09026981
CC GENERAL INFORMATION:
CC APPLICANT: Counter, Christopher M.
CC APPLICANT: Meyerson, Matthew
CC APPLICANT: Weinberg, Robert A.
CC TITLE OF INVENTION: Telomerase Catalytic Subunit Gene and
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Militia Drive
CC CITY: Lexington
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/026,981
CC FILING DATE: 20-FEB-1998
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/064,322
CC FILING DATE: 30-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/055,762
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/054,549
CC FILING DATE: 01-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/047,151
CC FILING DATE: 20-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/038,750
CC FILING DATE: 20-FEB-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WH197-11p4AM
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 781-861-6240
CC TELEFAX: 781-861-9540
CC INFORMATION FOR SEQ ID NO: 36:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 3346 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 3346 BP; 617 A; 1027 C; 965 G; 735 T; 2 OTHER.

Query Match 56.4%; Score 508; DB 23; Length 3346;
Best Local Similarity 79.8%; Pred. No. 0.00e+00;
Matches 719; Conservative 0; Mismatches 0; Indels 182; Gaps 1;
Db 1251 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGCTCGGGCCGATTGTGAACATGGACTA 1310
QY 1920 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGCTCGGGCCGATTGTGAACATGGACTA 1979
Db 1311 CGTCGTGGGAGCCAGAACGTTCCGACAGAAAAGAGGGCCGAGCGTCTCACTCGAGGGT 1370
QY 1980 CGTCGTGGGAGCCAGAACGTTCCGACAGAAAAGAGGGCCGAGCGTCTCACTCGAGGGT 2039
Db 1371 GAAGGCACCTGTTACGCTGCTCAACTACGAGCGGGCCGCGGCCCTCCCTGGGGGC 1430
QY 2040 GAAGGCACCTGTTACGCTGCTCAACTACGAGCGGGCCGCGGCCCTCCCTGGGGGC 2099
Db 1431 CTCTGTGCTGGGCTGACGATATCCACAGGGCTGCGGCACCTTCGTGTGCTGTGCG 1490
QY 2100 CTCTGTGCTGGGCTGACGATATCCACAGGGCTGCGGCACCTTCGTGTGCTGTGCG 2159
Db 1491 GGCCAGGACCCGCCCTGAGCTGTACTTTGTCAAGTGGATGTGACGGCGCGTACGA 1550
QY 2160 GGCCAGGACCCGCCCTGAGCTGTACTTTGTCAAGTGGATGTGACGGCGCGTACGA 2219
Db 1551 CACCATCCCCAGGACAGGCTCAGGAGGTATCGCCAGCATCATCAAAACCCAGAACAC 1610
QY 2220 CACCATCCCCAGGACAGGCTCAGGAGGTATCGCCAGCATCATCAAAACCCAGAACAC 2279
Db 1611 GTACTGCTGCTGGTATGCGGTATGCGGTGCTGAGAGGGCCGCGGCCATGGCAGGC 1670
QY 2280 GTACTGCTGCTGGTATGCGGTGCTGAGAGGGCCGCGGCCATGGCAGGC 2339
Db 1671 CTTCAAGAGCCACGTC----- 1686
QY 2340 CTTCAAGAGCCACGTCCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCTGTCG 2399
Db 1686 ----- 1686
QY 2400 TCACCTGCAGAGACACGACCCGCTGAGGGATGCCGTGCTCATCGAGCAGAGCTCCTCCCT 2459
Db 1686 ----- 1686
QY 2460 GAATGAGCCAGCAGTGGCTCTTCGAGGCTTCCTACGCTTCATGTGCCACACGCGGT 2519
Db 1687 -----CTACGTCCAGTGCAGGGGATCCCGCAGGGCTCCATCTCTC 1728
QY 2520 GCGCATCAGGGGCAAGTCCCTACGTCCAGTCCAGGGGATCCCGCAGGGCTCCATCTCTC 2579
Db 1729 CACGCTGCTGCAAGCCTGTGCTACGGCGACATGGAGAACAAAGCTGTTGGGGGATTCG 1788
QY 2580 CACGCTGCTGCAAGCCTGTGCTACGGCGACATGGAGAACAAAGCTGTTGGGGGATTCG 2639
Db 1789 GCGGAGCGGCTGCTGCTGGGTTGGTGGATGATTTCTTGTGTGACACCTCACCTCAC 1848
QY 2640 GCGGAGCGGCTGCTGCTGGGTTGGTGGATGATTTCTTGTGTGACACCTCACCTCAC 2699
Db 1849 CCACGCGAAAACCTTCCTCAGGACCTGGTCCGAGGTGTCCCTGAGTATGCTGCGTGGT 1908
QY 2700 CCACGCGAAAACCTTCCTCAGGACCTGGTCCGAGGTGTCCCTGAGTATGCTGCGTGGT 2759
Db 1909 GAACCTGGGGAAGACAGTGGTGAACCTTCCTCTGTAGAACAGAGGCCCTGGGTGACAGGC 1968
QY 2760 GAACCTGGGGAAGACAGTGGTGAACCTTCCTCTGTAGAACAGAGGCCCTGGGTGACAGGC 2819
Db 1969 T 1969
QY 2820 T 2820
RESULT 12
ID US-08-911-312-18 STANDARD; DNA; UNC; 3855 BP.
AC xxxxxx
DT

CC NUMBER OF SEQUENCES: 225
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002930US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 100:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2171 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 22..1716
CC OTHER INFORMATION: /note= "EcoRI-NotI insert of
CC OTHER INFORMATION: clone 712562 encoding 63 kDa
CC OTHER INFORMATION: telomerase protein"
CC SEQUENCE 2171 BP; 433 A; 567 C; 641 G; 429 T; 1 OTHER.

Query Match 52.9%; Score 477; DB 21; Length 2171;
Best Local Similarity 79.6%; Pred. No. 0.00e+00;
Matches 717; Conservative 1; Mismatches 0; Indels 183; Gaps 2;
Db 231 GTCCAGACTCGGCTTCATCCCAAGCTGACGGCTGGGCGGCGGATGTGAACATGGACTA 290
Qy 1920 GTCCAGACTCGGCTTCATCCCAAGCTGACGGCTGGGCGGCGGATGTGAACATGGACTA 1979
Db 291 CGTCGTGGGAGCCAGAGCTTCCGACAGAAAAAGAGGCCGCGGCTCTCACTCGAGGGT 350
Qy 1980 CGTCGTGGGAGCCAGAGCTTCCGACAGAAAAAGAGGCCGCGGCTCTCACTCGAGGGT 2039
Db 351 GAAGGCACTGTTCAGCGTGCTCAACTACGAGCGGCGCGGCGGCGGCTCTCTGGGCGC 410
Qy 2040 GAAGGCACTGTTCAGCGTGCTCAACTACGAGCGGCGGCGGCGGCGGCTCTCTGGGCGC 2099
Db 411 CTCTGTGCTGGGCGCTGACGATATCCACAGGGCGCTGGCGACCTTCGTGCTGCTGCG 470

Qy 2100 CTCTGTGCTGGGCGCTGGACGATATCCACAGGGCGCTGGCGACCTTCGTGCTGCTGCG 2159
Db 471 GGCCCAAGACCCCGCCCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGCGCGTACGA 530
Qy 2160 GGCCCAAGACCCCGCCCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGCGCGTACGA 2219
Db 531 CACCATCCCCCAGACAGGCTCAGGAGGTCTACGCCAGCATCATCAAAACCCAGACAC 590
Qy 2220 CACCATCCCCCAGACAGGCTCAGGAGGTCTACGCCAGCATCATCAAAACCCAGACAC 2279
Db 591 GTACTGCGTGGTGGTATGCCGTGGTCCAGAAAGGCCGCC-ATGGGGCACGTCGCGAAGGC 649
Qy 2280 GTACTGCGTGGTGGTATGCCGTGGTCCAGAAAGGCCGCCATGGGACGCTCCGCAAGGC 2339
Db 650 CTTCAAGAGCCACGTC-----
Qy 2340 CTTCAAGAGCCACGTCCTCTACTTTGACAGACCTCCAGCCGTACATGGACAGTTCGTTGC 2399
Db 665 -----
Qy 2400 TCACCTGCAGGAGACACAGCCCGCTGAGGGATGCCGTGTCATCGAGCAGAGCTCCTCCT 2459
Db 665 -----
Qy 2460 GAATGAGCCAGCAGTGGCGCTTTCGACGCTTCTCTAGGCTTCTATGTGCCACACGCCGT 2519
Db 666 -----CTACGTCCAGTCCAGGGGATCCCGCAGGGCTCCCATCTCTC 707
Qy 2520 GCGCATCAGGGGCAAGTCCTACGTCCAGTCCAGGGGATCCCGCAGGGCTCCCATCTCTC 2579
Db 708 CAGCTGCTCTGAGCCTGTGTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCG 767
Qy 2580 CAGCTGCTCTGAGCCTGTGTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCG 2639
Db 768 GCGGAGCGGCTGCTCTGCTGTTTGGTGGATGATTTCTTGTGTGACACTCACTCAC 827
Qy 2640 GCGGAGCGGCTGCTCTGCTGTTTGGTGGATGATTTCTTGTGTGACACTCACTCAC 2699
Db 828 CCACGCGAAAAACCTTCTCAGGACCCCTGTCGAGGTGTCCCTGAGTATGGTGCCTGGT 887
Qy 2700 CCACGCGAAAAACCTTCTCAGGACCCCTGTCGAGGTGTCCCTGAGTATGGTGCCTGGT 2759
Db 888 GAACCTCGGGAAGACAGTGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGCACGGC 947
Qy 2760 GAACCTCGGGAAGACAGTGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGCACGGC 2819
Db 948 T 948
Qy 2820 T 2820
Search completed: Thu Dec 24 19:19:19 1998
Job time : 2467 secs.

THIS PAGE BLANK (USPTO)

WORLD

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 18:29:00 1998; MasPar time 74.84 Seconds
Tabular output not generated. 561.595 Million cell updates/sec

Title: >US-08-951-733-13
Description: (1920-2820) from US08951733.seq (2 of 2)
Perfect Score: 901
N.A. Sequence: 1920 GTCCAGACTCGCTTCATCC.....AGGCCCTGGTGGCAGCGCT 2820
Comp: CAGCTCTGAGCGGAAGTAGG.....TCGGGACCCACCGTGCCGA

Scoring table: TABLE default
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 88822 seqs, 23323279 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-issued
1:5_COMB 2:PCT9_COMB 3:backfiles1

Statistics: Mean 8.416; Variance 4.601; scale 1.829

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | | | ID | Description | Pred. No. | |
|------------|-------|-------|--------|------|----------------------------------|----------------------------------|----------|
| | Score | Match | Length | | | | |
| 1 | 39 | 4.3 | 7218 | 1 | US-08-232-Sequence 14, Applicati | 4.38e-11 | |
| 2 | 30 | 3.3 | 215 | 1 | US-08-238-Sequence 5, Applicatio | 1.52e-05 | |
| C | 3 | 28 | 3.1 | 215 | US-08-238-Sequence 5, Applicatio | 2.24e-04 | |
| 4 | 26 | 2.9 | 74 | 2 | PCT-US95-1Sequence 94, Applicati | 3.08e-03 | |
| 5 | 26 | 2.9 | 81 | 2 | PCT-US95-1Sequence 92, Applicati | 3.08e-03 | |
| C | 6 | 25 | 2.8 | 74 | 2 | PCT-US95-1Sequence 100, Applicat | 1.11e-02 |
| 7 | 25 | 2.8 | 75 | 2 | PCT-US95-1Sequence 99, Applicati | 1.11e-02 | |
| C | 8 | 25 | 2.8 | 81 | 2 | PCT-US95-1Sequence 98, Applicati | 1.11e-02 |
| C | 9 | 25 | 2.8 | 81 | 2 | PCT-US95-1Sequence 92, Applicati | 1.11e-02 |
| 10 | 25 | 2.8 | 82 | 2 | PCT-US95-1Sequence 97, Applicati | 1.11e-02 | |
| C | 11 | 25 | 2.8 | 7218 | 1 | US-08-232-Sequence 14, Applicati | 1.11e-02 |
| 12 | 24 | 2.7 | 74 | 2 | PCT-US95-1Sequence 100, Applicat | 3.90e-02 | |
| C | 13 | 24 | 2.7 | 74 | 2 | PCT-US95-1Sequence 94, Applicati | 3.90e-02 |
| C | 14 | 24 | 2.7 | 75 | 2 | PCT-US95-1Sequence 99, Applicati | 3.90e-02 |
| C | 15 | 24 | 2.7 | 81 | 2 | PCT-US95-1Sequence 98, Applicati | 3.90e-02 |
| C | 16 | 24 | 2.7 | 82 | 2 | PCT-US95-1Sequence 97, Applicati | 3.90e-02 |
| 17 | 23 | 2.6 | 66 | 1 | US-08-471-Sequence 144, Applicat | 1.34e-01 | |
| C | 18 | 23 | 2.6 | 66 | 1 | US-08-471-Sequence 144, Applicat | 1.34e-01 |
| 19 | 23 | 2.6 | 69 | 1 | US-08-471-Sequence 142, Applicat | 1.34e-01 | |

| | | | | | | | |
|---|----|----|-----|------|---|----------------------------------|----------|
| c | 20 | 23 | 2.6 | 69 | 1 | US-08-471-Sequence 142, Applicat | 1.34e-01 |
| | 21 | 23 | 2.6 | 84 | 1 | US-08-209-Sequence 23, Applicati | 1.34e-01 |
| | 22 | 23 | 2.6 | 84 | 1 | US-08-133-Sequence 120, Applicat | 1.34e-01 |
| | 23 | 23 | 2.6 | 84 | 2 | PCT-US94-0Sequence 25, Applicati | 1.34e-01 |
| | 24 | 23 | 2.6 | 84 | 2 | PCT-US94-0Sequence 25, Applicati | 1.34e-01 |
| | 25 | 23 | 2.6 | 84 | 1 | US-08-300-Sequence 25, Applicati | 1.34e-01 |
| c | 26 | 23 | 2.6 | 242 | 1 | US-08-273-Sequence 1, Applicatio | 1.34e-01 |
| | 27 | 22 | 2.4 | 65 | 1 | US-08-471-Sequence 145, Applicat | 4.51e-01 |
| c | 28 | 22 | 2.4 | 68 | 1 | US-08-471-Sequence 143, Applicat | 4.51e-01 |
| c | 29 | 22 | 2.4 | 2116 | 1 | US-08-701-Sequence 1, Applicatio | 4.51e-01 |
| | 30 | 22 | 2.4 | 7001 | 1 | US-08-258-Sequence 1, Applicatio | 4.51e-01 |
| | 31 | 22 | 2.4 | 7001 | 1 | US-08-456-Sequence 1, Applicatio | 4.51e-01 |
| | 32 | 22 | 2.4 | 7001 | 1 | US-08-457-Sequence 1, Applicatio | 4.51e-01 |
| | 33 | 22 | 2.4 | 7001 | 1 | US-08-458-Sequence 1, Applicatio | 4.51e-01 |
| | 34 | 22 | 2.4 | 7001 | 1 | US-08-457-Sequence 1, Applicatio | 4.51e-01 |
| c | 35 | 21 | 2.3 | 59 | 2 | PCT-US95-1Sequence 95, Applicati | 1.47e+00 |
| | 36 | 21 | 2.3 | 65 | 1 | US-08-471-Sequence 145, Applicat | 1.47e+00 |
| c | 37 | 21 | 2.3 | 66 | 2 | PCT-US95-1Sequence 93, Applicati | 1.47e+00 |
| | 38 | 21 | 2.3 | 68 | 1 | US-08-471-Sequence 143, Applicat | 1.47e+00 |
| | 39 | 21 | 2.3 | 242 | 1 | US-08-273-Sequence 1, Applicatio | 1.47e+00 |
| | 40 | 21 | 2.3 | 957 | 1 | US-07-745-Sequence 16, Applicati | 1.47e+00 |
| | 41 | 21 | 2.3 | 1387 | 3 | 5270178-1Patent No. 5270178 | 1.47e+00 |
| | 42 | 21 | 2.3 | 1755 | 3 | 5225537-1Patent No. 5225537 | 1.47e+00 |
| | 43 | 21 | 2.3 | 5467 | 1 | US-07-745-Sequence 12, Applicati | 1.47e+00 |
| | 44 | 21 | 2.3 | 5904 | 2 | PCT-US94-0Sequence 1, Applicatio | 1.47e+00 |
| c | 45 | 20 | 2.2 | 2671 | 3 | 5168051-9Patent No. 5168051 | 4.67e+00 |

ALIGNMENTS

RESULT 1
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC xxxxxx
DE Sequence 14, Application US/08232463
CC Sequence 14, Application US/08232463
CC Patent No. 5670367
CC GENERAL INFORMATION:
CC APPLICANT: DORNER, F.
CC APPLICANT: SCHEIFLINGER, F.
CC APPLICANT: FALKNER, F. G.
CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22313-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/232,463
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/07/935,313
CC FILING DATE:
CC APPLICATION NUMBER: EP 91 114 300.6
CC FILING DATE: 26-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 30472/114 IMM
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703)836-9300
CC TELEFAX: (703)683-4109
CC TELEX: 899149
CC INFORMATION FOR SEQ ID NO: 14:

```
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: PT2gpt-F1s
SQ SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

Query Match 4.3%; Score 39; DB 1; Length 7218;
Best Local Similarity 1.1%; Pred. No. 4.38e-11;
Matches 4; Conservative 201; Mismatches 166; Indels 0; Gaps 0;

Db 1069 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1128
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 2354 TCTACCTTGACAGACCTCCAGCGTCATCGCAGACGTTCTGCTGCTCACCTGAGGAGA 2413
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1129 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1188
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 2414 CCAGCCCTGAGGATCGCTGTCATCGAGCAGCTCTCTCTGAATGAGCCAGCA 2473
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1189 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1248
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 2474 GTGGCTCTTCGACGCTTCTCTAGCTTCATGTCGCCACCGCGTGGCATAGGGGCA 2533
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1249 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1308
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 2534 AGTCCTAGCTCCAGTCGCGAGGGATCCGCGAGGCTCCATCTCTCCAGCTGCTGCA 2593
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1309 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1368
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 2594 CCTGTGTAGCGGACATGAGACACAGCTGTTCGGGGATCGCGGGACGGGTCG 2653
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1369 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1428
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 2654 TCCTGCTTGTGGTATGTTCTTGTGTGACACCTCACTCACCGCGGAAACCT 2713
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1429 YYYYYYGTACC 1439
: : : : : | | | |
Qy 2714 TCCTCAGGACC 2724

RESULT 2
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DE
DT
Sequence 5, Application US/08238163
CC Sequence 5, Application US/08238163
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: STOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourie and Crew
CC STREET: Steuart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
```

```
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC LOCATION: 1..215
CC NAME/KEY: misc.feature
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 3.3%; Score 30; DB 1; Length 215;
Best Local Similarity 14.8%; Pred. No. 1.52e-05;
Matches 16; Conservative 46; Mismatches 46; Indels 0; Gaps 0;

Db 1 MTNVTSSSVSTASCDKAKKDGNTTSWTFDCCNRTWGVCDTPTTYRVNNDGHNK 60
: : : : : | | : : : | : : : | : : : | : : : | : : :
Qy 2107 CTGGCGCTGACGATATCCACAGGCGCTGCGCACCTTCGTGCTGCTGCGGCGCCAG 2166
: : : : : | | : : : | : : : | : : : | : : : | : : :
Db 61 YSSANYGNGNNVCAKTHYTHNVSGADSKVTVDYSNASCSTSSSSNG 108
: : : : : | | : : : | : : : | : : : | : : : | : : :
Qy 2167 GACCCGCGCGCTGAGCTGTACTTTGTCAAGGTGATGTGACGGCGCG 2214
: : : : : | | : : : | : : : | : : : | : : : | : : :

RESULT 3
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DE
DT
Sequence 5, Application US/08238163
CC Sequence 5, Application US/08238163
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: STOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourie and Crew
CC STREET: Steuart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
```

```
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc_feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 3.1%; Score 28; DB 1; Length 215;
Best Local Similarity 14.1%; Pred. No. 2.24e-04;
Matches 24; Conservative 62; Mismatches 84; Indels 0; Gaps 0;

Db 35 DCCNRTGVCDDTTRVYVNDSSGHNKYSANYNGGNNVGAAKTHYTHYVNVGADSKTV 94
Cp 2506 ACATGACGCTAGGAGACGTCGAGAGGCCACTGCTGGCTCATTCAGGAGGAGCTCT 2447
Db 95 TDSYNASTSSNGTGDNRSGADSYGSKTAMTSRNTGKTANNVDSRNMGDASVGS 154
Cp 2446 GCTCATGACGACGCACTCCCTCAGCGGCTGCTCTCCTCGAGGTGAGCCAGCACTGTC 2387
Db 155 KNTKHKANSADGVGSKNNGDRNRYTGTKSVNSNCGGNGKRDVSSY 204
Cp 2386 GCATGTACGCGTGGAGGTCTCTCAAGGTAGAGACGTGGCTCTTGAAGGCC 2337

RESULT 4
ID PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DE
DT
DE Sequence 94, Application PC/TUS9511934
DE Sequence 94, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytozen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mirock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 94:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 74 BP; 3 A; 4 C; 3 G; 1 T; 63 OTHER.

SQ
```

```
Query Match 2.9%; Score 26; DB 2; Length 74;
Best Local Similarity 11.4%; Pred. No. 3.08e-03;
Matches 8; Conservative 19; Mismatches 43; Indels 0; Gaps 0;

Db 3 GAGNNNNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBN 62
Qy 2448 GAGCTCTCTCTGATGAGCCAGCAGTGGCTCTTCGACGCTTCTTACGCTTCA 2507
Db 63 BNNENACGCC 72
Qy 2508 CCACCACGCC 2517

RESULT 5
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DE
DT
DE Sequence 92, Application PC/TUS9511934
DE Sequence 92, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytozen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mirock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 92:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 81 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 81 BP; 3 A; 5 C; 6 G; 4 T; 63 OTHER.

SQ

Query Match 2.9%; Score 26; DB 2; Length 81;
Best Local Similarity 11.4%; Pred. No. 3.08e-03;
Matches 8; Conservative 19; Mismatches 43; Indels 0; Gaps 0;

Db 10 GAGNNNNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBN 69
Qy 2448 GAGCTCTCTCTGATGAGCCAGCAGTGGCTCTTCGACGCTTCTTACGCTTCA 2507
Db 70 BNNENACGCC 79
Qy 2508 CCACCACGCC 2517

RESULT 6
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
```


THIS PAGE BLANK (USPTO)

W P E R L H

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 19:23:48 1998; MasPar time 12290.10 Seconds
681.038 Million cell updates/sec

Tabular output not generated.

Title: >US-08-951-733-19
Description: (1-3798) from US08951733.seq
Perfect Score: 3798
N.A. Sequence: 1 CCACCGCTCCGGCGACGCT.....GGAATAGTCCATCCCTGAT 3798
Comp: GGTGGCGAGCGCCGTCGCA.....CCTTATCAGTAGGACGCA

Scoring table: TABLE default
Gap open 30; Gap extend 1
Nmatch STD : Dbase 0; Query 0

Searched: 567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb155
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_om 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_v1

Database: genbank107
15:gb_ba1 16:gb_ba2 17:gb_htg 18:gb_in 19:gb_om 20:gb_ov
21:gb_pat 22:gb_ph 23:gb_pl1 24:gb_pl2 25:gb_pr1
26:gb_pr2 27:gb_pr3 28:gb_ro 29:gb_st 30:gb_sts 31:gb_sy
32:gb_un 33:gb_v1

Statistics: Mean 11.964; Variance 6.048; scale 1.978

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|---------------------------------|-----------|
| 1 | 3784 | 99.6 | 4015 | 27 | AF015950 Homo sapiens telomeras | 0.00e+00 |
| 2 | 3770 | 99.3 | 4027 | 27 | AF018167 Homo sapiens telomeras | 0.00e+00 |
| 3 | 3439 | 90.5 | 8960 | 31 | AF043739 Synthetic construct hu | 0.00e+00 |
| 4 | 1079 | 28.4 | 3369 | 28 | AF073311 Mus musculus telomeras | 0.00e+00 |
| 5 | 1079 | 28.4 | 3426 | 28 | AF051911 Mus musculus telomeras | 0.00e+00 |
| 6 | 129 | 3.4 | 201 | 28 | AF029235 Mus musculus telomeras | 7.48e-73 |
| 7 | 67 | 1.8 | 7218 | 21 | I66494 Sequence 14 from patent | 5.33e-26 |
| 8 | 49 | 1.3 | 7218 | 21 | I66494 Sequence 14 from patent | 1.39e-13 |
| 9 | 37 | 1.0 | 74371 | 26 | AC005369 Homo sapiens chromosom | 5.37e-06 |
| 10 | 34 | 0.9 | 215 | 21 | I28278 Sequence 5 from patent | 3.10e-04 |
| 11 | 35 | 0.9 | 215 | 21 | I28278 Sequence 5 from patent | 8.17e-05 |
| 12 | 34 | 0.9 | 74371 | 26 | AC005369 Homo sapiens chromosom | 3.10e-04 |

| | | | | | | | |
|----|----|-----|--------|----|------------|------------------------|----------|
| 13 | 30 | 0.8 | 201 | 21 | A10161 | Synthetic DNA for prep | 5.30e-02 |
| 14 | 30 | 0.8 | 201 | 21 | A10158 | S.griseus gene for pre | 5.30e-02 |
| 15 | 30 | 0.8 | 201 | 21 | A10162 | Synthetic DNA for prep | 5.30e-02 |
| 16 | 30 | 0.8 | 201 | 21 | A10159 | S.griseus gene for pre | 5.30e-02 |
| 17 | 30 | 0.8 | 201 | 21 | E04076 | gDNA encoding envelope | 5.30e-02 |
| 18 | 29 | 0.8 | 6011 | 33 | HS4FSIR22T | Epstein-Barr virus (wi | 1.81e-01 |
| 19 | 29 | 0.8 | 15807 | 15 | MEU72662 | Methylobacterium extor | 1.81e-01 |
| 20 | 29 | 0.8 | 76094 | 25 | HS179D3A | Human DNA sequence fro | 1.81e-01 |
| 21 | 32 | 0.8 | 92879 | 17 | AC004450 | *** SEQUENCING IN PROG | 4.23e-03 |
| 22 | 31 | 0.8 | 216021 | 26 | HUAC004787 | Homo sapiens Chromosom | 1.51e-02 |
| 23 | 23 | 0.7 | 565 | 21 | E04076 | gDNA encoding envelope | 1.96e+00 |
| 24 | 27 | 0.7 | 904 | 33 | HS4BHLF | Epstein-Barr virus Bam | 1.96e+00 |
| 25 | 27 | 0.7 | 1572 | 19 | BTU21051 | Bos taurus DNA polymr | 1.96e+00 |
| 26 | 28 | 0.7 | 1663 | 19 | MVU92534 | Mustela vison microsat | 6.04e-01 |
| 27 | 26 | 0.7 | 2926 | 15 | SMPROBAG | S.marcescens proBA gen | 6.19e+00 |
| 28 | 27 | 0.7 | 4053 | 18 | DRO28DC162 | Drosophila melanogaste | 1.96e+00 |
| 29 | 27 | 0.7 | 5050 | 33 | HS4U2IR2A | Epstein-Barr virus (B9 | 1.96e+00 |
| 30 | 27 | 0.7 | 5400 | 15 | SCACTVA | S.coelicolor 6 actVA r | 1.96e+00 |
| 31 | 26 | 0.7 | 5591 | 15 | AF027439 | Acetobacter vinelandii | 6.19e+00 |
| 32 | 28 | 0.7 | 31812 | 24 | SPUNK4 | S.pombe chromosome I c | 6.04e-01 |
| 33 | 27 | 0.7 | 35980 | 16 | MTY20H10 | Mycobacterium tubercul | 1.96e+00 |
| 34 | 28 | 0.7 | 39073 | 23 | SPAC2E11 | S.pombe chromosome I c | 6.04e-01 |
| 35 | 26 | 0.7 | 43034 | 26 | HSN5H6 | Human DNA sequence fro | 6.19e+00 |
| 36 | 27 | 0.7 | 64078 | 18 | AC004365 | Drosophila melanogaste | 1.96e+00 |
| 37 | 27 | 0.7 | 81023 | 18 | AC001659 | Drosophila melanogaste | 1.96e+00 |
| 38 | 26 | 0.7 | 83957 | 27 | HS885E17 | Human DNA sequence *** | 1.96e+00 |
| 39 | 26 | 0.7 | 85132 | 26 | U82668 | Homo sapiens shox gene | 6.19e+00 |
| 40 | 27 | 0.7 | 92879 | 17 | AC004450 | *** SEQUENCING IN PROG | 1.96e+00 |
| 41 | 26 | 0.7 | 138706 | 17 | AC004946 | *** SEQUENCING IN PROG | 6.19e+00 |
| 42 | 26 | 0.7 | 147542 | 17 | HS754E20 | Human DNA sequence *** | 6.19e+00 |
| 43 | 27 | 0.7 | 172281 | 33 | EBV | Epstein-Barr virus (EB | 1.96e+00 |
| 44 | 27 | 0.7 | 184113 | 33 | HS4B958RAJ | Epstein-Barr virus, ar | 1.96e+00 |
| 45 | 27 | 0.7 | 216021 | 26 | HUAC004787 | Homo sapiens Chromosom | 1.96e+00 |

ALIGNMENTS

RESULT 1
LOCUS AF015950 4015 bp mRNA PRI 16-AUG-1997
DEFINITION Homo sapiens telomerase reverse transcriptase (hTERT) mRNA, complete cds.
ACCESSION AF015950
NID 92330016
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 4015)
AUTHORS Nakamura,T.M., Morin,G.B., Chapman,K.B., Weinrich,S.L., Andrews,W.H., Lingner,J., Harley,C.B. and Cech,T.R.
TITLE Telomerase catalytic subunit homologs from fission yeast and human
JOURNAL Science 277 (5328), 955-959 (1997)
MEDLINE 97400623
REFERENCE 2 (bases 1 to 4015)
AUTHORS Morin,G.B.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-1997) Geron Corporation, 230 Constitution Drive, Menlo Park, CA 94025, USA
FEATURES
source Location/Qualifiers
1..4015
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="kidney"
/map="5p"
/dev_stage="embryo"
/chromosome="5"
1..4015
/gene="hTERT"
56..3454
/gene="hTERT"
/codon_start=1

/product="telomerase reverse transcriptase"

/db_xref="pid:g3330017"

/translation="MPRAPRCRAVRSLRLSHRYEVLPLATFVRRLLPGQWRLLVQRGDP
AARFALVAOCLVCPWDAPPPAPSFQVSCILKELVARVLQRLGERGAKNVLAFGFA
LLDARGSGPEAFETVSRYLPTNTVDALRGSGAWGLLLRRVGGDVLVHLARCALFV
LVAPSCAYOVGPPLVOLGAATAORPPPHASGPRRLRGCEERANWHSVREAGVPLGLPA
PGARRGGSASRSLPLPKPRPRGAAPERTPVQGSWAHPGRTGRSPDRGRVYSPFA
RPAEATISLEGALSGTRNHSVSRQHHAGPSTSRPPRWDTPCPFPVIAETKHFLLYS
SGDKEQLRFSFLUSSRLPSLTGARLVETIFLGRSPMFGTPRRLPLPQRYWOMRPL
FLLEGNHQAQCTPGVLLKTHCPLRAAETPAAGVCAREKPGQSVAAPEEDTDPRLVQ
LLRQHSFMQVYFVRACLRLRPVPLGLWGRNRRERFLNKKFISLGHKARSLQEL
TKMSVRDCAWLRRSPGVGCPAAEHLREEILAKFLHMLSVVYVELLRSFYVET
TFQKNRLFYRKSWNSKLQSIGIRQHLKRVQLSELSEAEVRQHRARPALLLSRLAFI
PFQDGLRPIVNMIDYVGAQTFRREKRAERLTSRKALFVSLNTERARFGLLGASVLG
LDDIHWARTFLVRQAQDPPPELFYKVDVTGAYDTIPQDRLTEYIAIIIPONTY
VRYAVVQAHHGVKRAKFSKSHVSTLDLOPMNRQFVAHQETSPILRDVAVTEQSSSL
NEASSGLFDVLFEMCHAVIRKSKYVOCQPOGSLSTLCLSCYGDMEKNLFA
IRRDGLLLRVDLFLVTLPHLTHAKTFLRLVRGVEYGCNVNLRKTVNFPVEDEAL
GGTAFVQMPAHGLFPKGLLLDTRILEVQSDYSSYARTSIRASLTNRGFKAGRNWR
KLFGLVRLKHSGLFDLQVNSLQTVCTNIYKILLQYRHFACVLQLPFHQQVWKNPT
FLFURVSDTASLCSILKAKNAGMSLGAAGPLPSEAVQWILCHOAFLKLTRHRT
YVPLGLSLTAQQLSRKLPGLTTLTALAANAALPALSDFKTLTD"

BASE COUNT 663 a 1363 c 1275 g 714 t

Query Match 99.6%; Score 3784; DB 27; Length 4015;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 3785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| Db | 1 | GCAGCGCTGGTCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 60 |
| QY | 13 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 72 |
| Db | 61 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 120 |
| QY | 73 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 132 |
| Db | 121 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 180 |
| QY | 133 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 192 |
| Db | 181 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 240 |
| QY | 193 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 252 |
| Db | 241 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 300 |
| QY | 253 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 312 |
| Db | 301 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 360 |
| QY | 313 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 372 |
| Db | 361 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 420 |
| QY | 373 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 432 |
| Db | 421 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 480 |
| QY | 433 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 492 |
| Db | 481 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 540 |
| QY | 493 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 552 |
| Db | 541 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 600 |
| QY | 553 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 612 |
| Db | 601 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 660 |
| QY | 613 | GCAGCGCTGGCTCTGTCGCGACGTGGAGCCCTGGCCCGCCGACCCCGCGATGCC | 672 |

| | | | |
|----|------|--|------|
| Db | 661 | CTGGAACCATAGCGTCAGGAGGCCGGGGTCCCGCTGGCGCTGCCAGCCCCCGGGTGGAG | 720 |
| QY | 673 | CTGGAACCATAGCGTCAGGAGGCCGGGGTCCCGCTGGCGCTGCCAGCCCCCGGGTGGAG | 732 |
| Db | 721 | GAGCGCGGGGGCAGTGCAGCGGAGTCTGCCGTTGCCCAAGAGCCCGAGGCGTGGCG | 780 |
| QY | 733 | GAGCGCGGGGGCAGTGCAGCGGAGTCTGCCGTTGCCCAAGAGCCCGAGGCGTGGCG | 792 |
| Db | 781 | TGCCCCCTGAGCCGAGCGGAGCGGCGGCTTGGCGAGGGGTCTTGGCGGCGGCGAGGAC | 840 |
| QY | 793 | TGCCCCCTGAGCGGAGCGGAGCGGCGGCTTGGCGAGGGGTCTTGGCGGCGGCGAGGAC | 852 |
| Db | 841 | GCCTGAGCGAGTGACCGTGGTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 900 |
| QY | 853 | GCCTGAGCGAGTGACCGTGGTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 912 |
| Db | 901 | CACCTCTTTGGAGGGTGGCTCTCTGCGAGCGGCGGCTTGTGTGTGTGTGTGTGTGTGTGT | 960 |
| QY | 913 | CACCTCTTTGGAGGGTGGCTCTCTGCGAGCGGCGGCTTGTGTGTGTGTGTGTGTGTGTGT | 972 |
| Db | 961 | GCACACGCGGGGCGGCGGCGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 1020 |
| QY | 973 | GCACACGCGGGGCGGCGGCGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 1032 |
| Db | 1021 | CCCGGTGTAGCGCGAGACCAAGACATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1080 |
| QY | 1033 | CCCGGTGTAGCGCGAGACCAAGACATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1092 |
| Db | 1081 | GCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1140 |
| QY | 1093 | GCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1152 |
| Db | 1141 | GACCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1200 |
| QY | 1153 | GACCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1212 |
| Db | 1201 | GCCCCAGCGCTACTTGCGAAATGCGGCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1260 |
| QY | 1213 | GCCCCAGCGCTACTTGCGAAATGCGGCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1272 |
| Db | 1261 | GTGCGCTTCTGCGGGTCTCTCAAGACGACATGCGCGCTGCGAGCTGCGGCTTCTTCTTCTTCT | 1320 |
| QY | 1273 | GTGCGCTTCTGCGGGTCTCTCTCAAGACGACATGCGCGCTGCGAGCTGCGGCTTCTTCTTCTTCT | 1332 |
| Db | 1321 | AGCGGCTGTCTGTGCGCGGAGAGCCCGAGGGCTCTGTGCGCGGCCCGCGAGGAGGAGGA | 1380 |
| QY | 1333 | AGCGGCTGTCTGTGCGCGGAGAGCCCGAGGGCTCTGTGCGCGGCCCGCGAGGAGGAGGA | 1392 |
| Db | 1381 | CACAGACCCCGCTGCGCTGGT | 1440 |
| QY | 1393 | CACAGACCCCGCTGCGCTGGT | 1452 |
| Db | 1441 | CGGCTTCTGCGGGCTGCTGCGCGGGTGTGCGGCCCGCGAGGCTTCTTCTTCTTCTTCTTCTTCT | 1500 |
| QY | 1453 | CGGCTTCTGCGGGCTGCTGCGCGGGTGTGCGGCCCGCGAGGCTTCTTCTTCTTCTTCTTCTTCT | 1512 |
| Db | 1501 | CAACGAACCGCGCTTCTTCTGAGGAACCAAGAGTTCATCTTCTTCTTCTTCTTCTTCTTCTTCT | 1560 |
| QY | 1513 | CAACGAACCGCGCTTCTTCTGAGGAACCAAGAGTTCATCTTCTTCTTCTTCTTCTTCTTCTTCT | 1572 |
| Db | 1561 | GCTCTCGCTGACGAGCTGACGTGGAAGATGAGCTGCGGGAGTGTGCTTCTTCTTCTTCTTCTTCT | 1620 |
| QY | 1573 | GCTCTCGCTGACGAGCTGACGTGGAAGATGAGCTGCGGGAGTGTGCTTCTTCTTCTTCTTCTTCT | 1632 |
| Db | 1621 | GAGCCAGGGGTGGCTGTGTTCCGCGCGCGAGACCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1680 |
| QY | 1633 | GAGCCAGGGGTGGCTGTGTTCCGCGCGCGAGACCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1692 |
| Db | 1681 | CAAGTTCCTGCACTGGCTGATGAGTGTACGTGCTGAGCTGCTCAGGCTCTTCTTCTTCTTCTTCTTCT | 1740 |
| QY | 1693 | CAAGTTCCTGCACTGGCTGATGAGTGTACGTGCTGAGCTGCTCAGGCTCTTCTTCTTCTTCTTCTTCT | 1752 |
| Db | 1741 | TGTCAGGAGACCAACCGCTTTCAAAGAACAGGCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1800 |

1753 TGTACGAGACACGCTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAG 1812
1801 CAAAGTTGCAAGCACTTGAATCAGACAGCACTTGAAGAGGTTGACAGCTGCGGGAGCTGTC 1860
1813 CAAAGTTGCAAGCACTTGAATCAGACAGCACTTGAAGAGGTTGACAGCTGCGGGAGCTGTC 1872
1861 GGAAGCAGAGTTCAGGACAGATCGGGAAGCAGAGCCCGCCCTGCTGAGTCTCAGACTCGG 1920
1873 GGAAGCAGAGTTCAGGACAGATCGGGAAGCAGAGCCCGCCCTGCTGAGTCTCAGACTCGG 1932
1921 CTTTATCCCCAAGCTGACGGCTCGGCCGAGTTCTGAACATGGACTAGCTGCTGGGAGC 1980
1933 CTTTATCCCCAAGCTGACGGCTCGGCCGAGTTCTGAACATGGACTAGCTGCTGGGAGC 1992
1981 CAGAACGTTCCGACAGAGAAAGAGGCGGAGCGTCTCACTCGAGGGTGAAGGCACTGTT 2040
1993 CAGAACGTTCCGACAGAGAAAGAGGCGGAGCGTCTCACTCGAGGGTGAAGGCACTGTT 2052
2041 CAGCGTGTCTCAACTACGAGGCGGCGGCGCCCGCCCTGCTGGGCGCTCTGTGCTGGG 2100
2053 CAGCGTGTCTCAACTACGAGGCGGCGGCGCCCGCCCTGCTGGGCGCTCTGTGCTGGG 2112
2101 CCGTGACGATATCCACAGGGCTGCGGCACCTTCTGCTGCTGGTGTGCGGGCCAGGACCC 2160
2113 CCGTGACGATATCCACAGGGCTGCGGCACCTTCTGCTGCTGGTGTGCGGGCCAGGACCC 2172
2161 GCCCGCTGAGCTGTACTTTGTTCAAGGTGGATGTGACGGCGCGGTACGACACCATCCCCCA 2220
2173 GCCCGCTGAGCTGTACTTTGTTCAAGGTGGATGTGACGGCGCGGTACGACACCATCCCCCA 2232
2221 GGACAGGCTCAGGAGGTATCGCCAGAGCATATCAAAACCCAGAACACAGTACTGCGTGG 2280
2233 GGACAGGCTCAGGAGGTATCGCCAGAGCATATCAAAACCCAGAACACAGTACTGCGTGG 2292
2281 TCGGTATGCGGTGTCAGAGGCGCCCATGGGACGCTTCCGCAAGGCTTCAAGAGCCA 2340
2293 TCGGTATGCGGTGTCAGAGGCGCCCATGGGACGCTTCCGCAAGGCTTCAAGAGCCA 2352
2341 CGTCTCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTTCGTGGCTCACTTCAGGA 2400
2353 CGTCTCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTTCGTGGCTCACTTCAGGA 2412
2401 GACACGCGCTGAGGAGTACGCGTGTATCGAGCAGAGTCTCTCCCTGGAATGAGGCGAG 2460
2413 GACACGCGCTGAGGAGTACGCGTGTATCGAGCAGAGTCTCTCCCTGGAATGAGGCGAG 2472
2461 CAGTGGCTCTTCGACGCTTCTTACGCTTTCATGTGCGCACACGCGCTGCGCATCAGGG 2520
2473 CAGTGGCTCTTCGACGCTTCTTACGCTTTCATGTGCGCACACGCGCTGCGCATCAGGG 2532
2521 CAAAGTCTACGTCAGTGCAGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2580
2533 CAAAGTCTACGTCAGTGCAGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2592
2581 CAGCCTGTGCTACGCGGACATGGAGAACAGCTGTTTGGGGGATTCGCGGGAGCGGCT 2640
2593 CAGCCTGTGCTACGCGGACATGGAGAACAGCTGTTTGGGGGATTCGCGGGAGCGGCT 2652
2641 GCTCTCGCTTGGTGGATGATTTCTTGTGTGACACTCACTCACTCACCCAGCGGAAAC 2700
2653 GCTCTCGCTTGGTGGATGATTTCTTGTGTGACACTCACTCACTCACCCAGCGGAAAC 2712
2701 CTTCTCTCAGGACCTTGGTCCGAGGTGCTCCTGAGTATGGCTGCGTGGTGAATTCGCGAA 2760
2713 CTTCTCTCAGGACCTTGGTCCGAGGTGCTCCTGAGTATGGCTGCGTGGTGAATTCGCGAA 2772
2761 GACAGTGTGAATCTCCTGTAGAGACAGGCGCTTGGGTGGCAGCGGCTTTGTTTCAGAT 2820
2773 GACAGTGTGAATCTCCTGTAGAGACAGGCGCTTGGGTGGCAGCGGCTTTGTTTCAGAT 2832
2821 GCCGGCCACGGCCCTATTCCCTCGTGGGCGCTGCTGTGTGATATCCCGGACCCCTGAGGT 2880

2833 GCCGCCACAGCGCTATTCCCTGGTGGGGCTGCTGCTGGATACCCGGAGCCCTGGAGGT 2892
Db CAGAGCGACTACTCCAGCTATGCCCGACCTCCATCAGAGCCAGTCTCACTTCAACCG 2940
2893 GCAGAGCGACTACTCCAGCTATGCCCGACCTCCATCAGAGCCAGTCTCACTTCAACCG 2952
Db CGGCTTCAAGGCTGGGAGGAACATCGCTCGCAAACTCTTTGGGGTCTTTGGCGCTGAAGTG 3000
Qy CGGCTTCAAGGCTGGGAGGAACATCGCTCGCAAACTCTTTGGGGTCTTTGGCGCTGAAGTG 3012
Db TCAGAGCGCTTTCTGGATTTGCAAGTGAACAGCTCCAGAGCGTGTGCACCAACATCTA 3060
Qy TCAGAGCGCTTTCTGGATTTGCAAGTGAACAGCTCCAGAGCGTGTGCACCAACATCTA 3072
Db CAAAGATCCTCTGCTGACAGCGTACAGGTTTACGCAATGTGTGCTGCAAGTCTCCATTTCA 3120
Qy CAAAGATCCTCTGCTGACAGCGTACAGGTTTACGCAATGTGTGCTGCAAGTCTCCATTTCA 3132
Db TCAGCAAGTTTGGAGAACCCACATTTTCTGCGGCTCATCTCTGACAGCGCTCCCT 3180
Qy TCAGCAAGTTTGGAGAACCCACATTTTCTGCGGCTCATCTCTGACAGCGCTCCCT 3192
Db CTGCTACTTCCATCTGAAAGCCAAAGCGCAGGATGTGCTGGGGCCAAAGGGCGCGC 3240
Qy CTGCTACTTCCATCTGAAAGCCAAAGCGCAGGATGTGCTGGGGCCAAAGGGCGCGC 3252
Db CGGCCCTCTGCCCTCCGAGGCGGTGCAAGTGTGTGCCAACCAAGCATTTCTGCTCAAGCT 3300
Qy CGGCCCTCTGCCCTCCGAGGCGGTGCAAGTGTGTGCCAACCAAGCATTTCTGCTCAAGCT 3312
Db GACTCGACACCGTGTACCTAGCTGCCACTCTGCGGCTCATCTCTGACAGCGCTCCCT 3360
Qy GACTCGACACCGTGTACCTAGCTGCCACTCTGCGGCTCATCTCTGACAGCGCCAGCGCA 3372
Db GCTGAGTCGGAAGCTCCCGGGGACGAGCTGCTGAGTGGCGCCAGGCGCCAGCCAGCGC 3420
Qy GCTGAGTCGGAAGCTCCCGGGGACGAGCTGCTGAGTGGCGCCAGGCGCCAGCCAGCGC 3432
Db ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGATGGCACCCGCCACAGCCAGGCGCA 3480
Qy ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGATGGCACCCGCCACAGCCAGGCGCA 3492
Db GAGCAGACACAGAGCGCTCTACGCGGGCTCTACCTCCAGGAGGAGGCGCGGCC 3540
Qy GAGCAGACACAGAGCGCTCTACGCGGGCTCTACCTCCAGGAGGAGGCGCGGCC 3552
Db CACACCCAGGCGCCAGCGCTGGGAGTCTGAGGCTGAGTGGTGGCGGAGGCGCTG 3600
Qy CACACCCAGGCGCCAGCGCTGGGAGTCTGAGGCTGAGTGGTGGCGGAGGCGCTG 3612
Db CATGTCCGGCTGAAGGCTGAGTGTCCGGCTCAGGCGCTGAGGAGTGTCCAGCCAGGCGT 3660
Qy CATGTCCGGCTGAAGGCTGAGTGTCCGGCTCAGGCGCTGAGGAGTGTCCAGCCAGGCGT 3672
Db GAGTGTCCAGCAGACCTGCGCTCTTCACTTCCACAGGCTGCGGCTCGGCTCCACCCCA 3720
Qy GAGTGTCCAGCAGACCTGCGCTCTTCACTTCCACAGGCTGCGGCTCGGCTCCACCCCA 3732
Db GGGCCAGCTTTTCTCACCAGGAGCGCGGCTTCCACTCCCCACATAGGAATAGTCCATCC 3780
Qy GGGCCAGCTTTTCTCACCAGGAGCGCGGCTTCCACTCCCCACATAGGAATAGTCCATCC 3792
Db CCAGAT 3786
Qy CCTGAT 3798

RESULT 2
LOCUS AF018167 4027 bp mRNA PRI 29-AUG-1997
DEFINITION Homo sapiens telomerase catalytic subunit (hST2) mRNA, complete cds.
ACCESSION AF018167
NID g2347128

```

KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea;
            Homo.

REFERENCE
AUTHORS      Meyerson, M., Counter, C.M., Eaton, E.N., Ellisen, L.W., Steiner, P.,
            Caddle, S.D., Ziaugra, L., Beijersbergen, R.L., Davidoff, M.J., Liu, Q.,
            Bacchetti, S., Haber, D.A., and Weinberg, R.A.
TITLE        hEST2, the putative human telomerase catalytic subunit gene, is
            up-regulated in tumor cells and during immortalization
JOURNAL      Cell 90 (4), 785-795 (1997)
MEDLINE      97433088
AUTHORS      1 (bases 1 to 4027)
            2 (bases 1 to 4027)
            Meyerson, M., Counter, C.M., Eaton, E.N., Ellisen, L.W., Steiner, P.,
            Caddle, S.D., Ziaugra, L., Beijersbergen, R.L., Davidoff, M.J., Liu, Q.,
            Bacchetti, S., Haber, D.A., and Weinberg, R.A.
TITLE        Direct Submission
JOURNAL      Submitted (11-AUG-1997) Whitehead Institute for Biomedical
            Research, Massachusetts Institute of Technology, Nine Cambridge
            Center, Cambridge, MA 02142, USA
FEATURES
Source      1..4027
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="5"
            /map="5p15.33"
            1..4027
            /note="human Ever Shorter Telomeres 2"
            /gene="hEST2"
            /gene="hEST2"
            /note="member of reverse transcriptase class"
            /codon_start=1
            /product="telomerase catalytic subunit"
            /db_xref="PID:92347129"
            /translation="MPRAPRCRAVRSLLRSHYREVLPATFVRRLPGOGWRLVORGDP
            AAFRALVAQVLCVFPAPPPAPSPFQVSLKELVARLQRCERGAKNVLAFGFA
            LIDGARGCPPEAFATSVRSYLPNTVTDALRSGAWGLLRVGDVLLHARLALFV
            LVAPSCATQVCPPIYOLGAATQARPPIHASPRLRGCERANWHSVREAGVPLGLPA
            PARRRRGASRLPLPKPRRGAPEPTPVGQSWAHPGRTGPRSDRGCFVVSVA
            PPAEATSLGALSTRHSVPSVGRQHAGPPSTRPWPDTFCPPVYATKHELYS
            SDKQLRPSFLLSRLSLTGARLVETIFLGRPMPPTPRRLPQRPYIWMRPL
            FLELGNHAQCPYGLLTKHCPRLAAVTPAAGCAAREKPOGSAPEEEDDPRLVQ
            LLRQSSPQVYGFYACRLRLPLPGLMGRSHNERRLTKFTLSLGHAKLSLQEL
            TWKMSVRCGATRRSPGVGCPAAEHLRREILAKFLHLSVYVVELRFFVYVET
            TFQKRLFYRKSVMKLSIGIRQHLKRVOLRSEAEVRQREARPAALITSRLFI
            PKPDGLRPIVNDYVVGARTFRREKAEALRSYKALFSVLNYEARPRPGLLGSYLG
            LDDIHRWRTFVLRVRAQDPPELIFVKVDVITGAYDTIPQDRLEIVIASILKPTQYIC
            VRYAVVQKAHGHVRKAFKSHVSTLTDLPYMRQFVAHQLETSPLRDVAVIEQSSSL
            NEASSGLDFVLRMCHHVRIRGKSYVQCQIGPOGSIILSLILSLCYGDMENKLFAG
            IRRDGLLRLLVDLFLVPHLTHTAKTFLTRVGRPEYGVNLRKTVVNFVEDEAL
            GGTAFVQWPAHGLFPMCGLLDTRTELVQSDYSVARTSIRASLTFRNGFRAGNMR
            KLFGRVLRKCHSLFDLQVNSLQVCTNIYKILLQAYRFHACVLQLPFHQQYWNRP
            FFLRVISDTSALCYSLIKAKNAGMSLGAGAPLPSPSEAVOMLCHQAFLLKLRHVT
            YVPLLSLRTAQLSRPLPTTLTALFAAANPALPSPDKILD"
BASE COUNT  674 a 1361 c 1277 g 715 t
ORIGIN
Query Match      99.3%; Score 3770; DB 27; Length 4027;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 3773; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 12  GTCTCTGCGGACGTGGGAAGCCCTGGCCCGCCGACCCCGCGATGCGCGCGCTCCC 71
      |||
Qy 23  GTCTCTGCGGACGTGGGAAGCCCTGGCCCGCCGACCCCGCGATGCGCGCGCTCCC 82
      |||
Db 72  CGTGTCCGAGCCGTGGCGCTCCCTGTGCGCAGCCACTACCGGAGGTGCTGCCGTGGCC 131
      |||
Qy 83  CGTGTCCGAGCCGTGGCGCTCCCTGTGCGCAGCCACTACCGGAGGTGCTGCCGTGGCC 142
      |||
Db 132 ACGTTGCTGGCGGCTTGGGGCCCGGAGGCTGGCGGCTGGTGCAGCGCGGGGACCGCGG 191
      |||

```

| | | | |
|----|------|--|------|
| Qy | 1223 | TACTGGCAATGCGGCCCTGTTTCTTCTGGAGCTGCTTGGGAACACAGCGCAGTGC0CCCTAC | 1288 |
| Db | 1272 | GGGCTGCTCTCAAGACGCACTCCCGCTGGCAGCTGCGGTCA0CC0CAGACGCGGTGC | 1331 |
| | | | |
| | | | |
| Qy | 1283 | GGGCTGCTCTCAAGACGCACTGCCCGTGGCAGCTGCGGTCA0CC0CAGACGCGGTGC | 1342 |
| Db | 1332 | TGTGCCGGGAGAAGCC0CAGGGCTGTGTGGCGGCC0CCGAGGAGGACACAGACCCC | 1391 |
| | | | |
| | | | |
| Qy | 1343 | TGTGCCGGGAGAAGCC0CAGGGCTGTGTGGCGGCC0CCGAGGAGGACACAGACCCC | 1402 |
| Db | 1392 | CGTGCCTGTGTGAGCTGCTCCGCCAGCACACAGACGCCCTCGCAGGTGTACGGCTCGTG | 1451 |
| | | | |
| | | | |
| Qy | 1403 | CGTGCCTGTGTGAGCTGCTCCGCCAGCACACAGACGCCCTGGCAGGTGTACGGCTCGTG | 1462 |
| Db | 1452 | CGGGCTTGCTGCGCGCGCTGTGCCCCAGGCCCTCTGGGCTCTCAGSCACACAGAACGC | 1511 |
| | | | |
| | | | |
| Qy | 1463 | CGGGCTTGCTGCGCGCGCTGTGCCCCAGGCCCTCTGGGCTCTCAGSCACACAGAACGC | 1522 |
| Db | 1512 | CGTCTCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAGCATGCCAGCTCTCGCTG | 1571 |
| | | | |
| | | | |
| Qy | 1523 | CGTCTCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAGCATGCCAGCTCTCGCTG | 1582 |
| Db | 1572 | CAGGACTGACGTGGAAGATGAGCTGCGGGGCTGCGCTTGCTGCGCAGGAGCCACAGG | 1631 |
| | | | |
| | | | |
| Qy | 1583 | CAGGACTGACGTGGAAGATGAGCTGCGGGGACTCGCTTGCTGCGCAGGAGCC0CAGG | 1642 |
| Db | 1632 | GTGTGGCTGTGTTCCGGCC0CAGACACCGCTCTGCTGAGGAGATCCTTGCCCAAGTTCCTG | 1691 |
| | | | |
| | | | |
| Qy | 1643 | GTGTGGCTGTGTTCCGGCC0CAGACACCGCTCTGCTGAGGAGATCCTTGCCCAAGTTCCTG | 1702 |
| Db | 1692 | CAC TGCTGATGAGTGTACGTCTGCGAGCTGCTCAGGTCTTTCTTTTATGTACCGGAG | 1751 |
| | | | |
| | | | |
| Qy | 1703 | CAC TGCTGATGAGTGTACGTCTGCGAGCTGCTCAGGTCTTTCTTTTATGTACCGGAG | 1762 |
| Db | 1752 | ACCACGTTTCAAAGAAACAGGCTCTTTTCTACCCGAAGAGTGTCTGGAGCAAGTGC0AA | 1811 |
| | | | |
| | | | |
| Qy | 1763 | ACCACGTTTCAAAGAAACAGGCTCTTTTCTACCCGAAGAGTGTCTGGAGCAAGTGC0AA | 1822 |
| Db | 1812 | AGCATTGAATCAGACAGCACTTGAAGAGGTGTACGTCTCGGAGACTGTCTGGAAAGCAGAG | 1871 |
| | | | |
| | | | |
| Qy | 1823 | AGCATTGAATCAGACAGCACTTGAAGAGGTGTACGTCTCGGAGACTGTCTGGAAAGCAGAG | 1882 |
| Db | 1872 | GTCAGGCAGCATCGGGA0CCAGG0CC0CCTGTGTGACGTCCAGACTCC0GTTCAATCCC | 1931 |
| | | | |
| | | | |
| Qy | 1883 | GTCAGGCAGCATCGGGA0CCAGG0CC0CCTGTGTGACGTCCAGACTCC0GTTCAATCCC | 1942 |
| Db | 1932 | AAGCCTGACGGGCTGCGGCCGATTTGTAACATTGTGAACTAGTCTGCGGAGCCAGAACGTT | 1991 |
| | | | |
| | | | |
| Qy | 1943 | AAGCCTGACGGGCTGCGGCCGATTTGTAACATTGTGAACTAGTCTGCGGAGCCAGAACGTT | 2002 |
| Db | 1992 | CGCAGAGAAAAGAGGCGCAGGCTCTACCTCGAGGGTGAAGGCACATGTTTCAGGCTGCTC | 2051 |
| | | | |
| | | | |
| Qy | 2003 | CGCAGAGAAAAGAGGCGCAGGCTCTACCTCGAGGGTGAAGGCACATGTTTCAGGCTGCTC | 2062 |
| Db | 2052 | AAC TACAGCGGCGCGCGCC0C0C0CCTCTCTGGCGCTCTGTGCTGGGCTGAGACGAT | 2111 |
| | | | |
| | | | |
| Qy | 2063 | AAC TACAGCGGCGCGCGCC0C0C0CCTCTCTGGCGCTCTGTGCTGGGCTGAGACGAT | 2122 |
| Db | 2112 | ATCCACAGGGCCTGGCGCACCTTCTGTCTGTGCTGTGGGGCC0CAGGACCC0C0CCTGAG | 2171 |
| | | | |
| | | | |
| Qy | 2123 | ATCCACAGGGCCTGGCGCACCTTCTGTCTGTGCTGTGGGGCC0CAGGACCC0C0CCTGAG | 2182 |
| Db | 2172 | CTGTACTTTGTCAAGTGTGATGTGACGGGCGGTACGACACCATCTCC0CAGGACAGGCTC | 2231 |
| | | | |
| | | | |
| Qy | 2183 | CTGTACTTTGTCAAGTGTGATGTGACGGGCGGTACGACACCATCTCC0CAGGACAGGCTC | 2242 |
| Db | 2232 | ACGAGGATCATCGCCAGCATCATCAACCC0CAGAACAGCTACTGCGTGGCTGGATGCC | 2291 |
| | | | |
| | | | |
| Qy | 2243 | ACGAGGATCATCGCCAGCATCATCAACCC0CAGAACAGCTACTGCGTGGCTGGATGCC | 2302 |
| Db | 2292 | GTGTTCAGAGGCGGCC0C0C0C0CAGTCTGCCAAGGCTCTCAAGAGCCAGCTCTCTACC | 2351 |
| | | | |
| | | | |
| Qy | 2303 | GTGTTCAGAGGCGGCC0C0C0C0CAGTCTGCCAAGGCTCTCAAGAGCCAGCTCTCTACC | 2362 |

| | | | |
|----|------|---|------|
| Db | 2352 | TTGACAGACCTCCAGCCGTACATGCGACAGTTCTGTGGCTCACTCTCAGGAGACCAAGCCCG | 2411 |
| Qy | 2363 | TTGACAGACCTCCAGCCGTACATGCGACAGTTCTGTGGCTCACTCTCAGGAGACCAAGCCCG | 2422 |
| Db | 2412 | CTGAGGGATCCGTCGTATCATGACAGAGACTCCTCCCTGAATGAGGCCAGCAGTGGCCCTC | 2471 |
| Qy | 2423 | CTGAGGGATCCGTCGTATCATGACAGAGACTCCTCCCTGAATGAGGCCAGCAGTGGCCCTC | 2482 |
| Db | 2472 | TTCGACGCTCTCCTACGCTTTCATGTGCCACACACGCGCTGCGCATCAGGGGCAAGTCCCTAC | 2531 |
| Qy | 2483 | TTCGACGCTCTCCTACGCTTTCATGTGCCACACGCGCTGCGCATCAGGGGCAAGTCCCTAC | 2542 |
| Db | 2532 | GTCCAGTGCCAGGGATCCCGCAGGGCTCCATCTCTCCACGCTCCTCTGCAGCCTGTGC | 2591 |
| Qy | 2543 | GTCCAGTGCCAGGGATCCCGCAGGGCTCCATCTCTCCACGCTCCTCTGCAGCCTGTGC | 2602 |
| Db | 2592 | TACGGCGACATGAGAACAAAGCTGTTTGGGGGATTCGGCGGAGCGGCTGCTCTGGCT | 2651 |
| Qy | 2603 | TACGGCGACATGAGAACAAAGCTGTTTGGGGGATTCGGCGGAGCGGCTGCTCTGGCT | 2662 |
| Db | 2652 | TTGCTGGATGATTCTTGTGTGTGACACCTCACTCAACCACGCGAAACCTTCTCTCAGG | 2711 |
| Qy | 2663 | TTGCTGGATGATTCTTGTGTGTGACACCTCACTCAACCACGCGAAACCTTCTCTCAGG | 2722 |
| Db | 2712 | ACCTGTGTCGAGGTGTCCCTGAGTATGGCTGCGTGGTGAATTCGCGAAGACAGTGGTG | 2771 |
| Qy | 2723 | ACCTGTGTCGAGGTGTCCCTGAGTATGGCTGCGTGGTGAATTCGCGAAGACAGTGGTG | 2782 |
| Db | 2772 | AACHTCCCTGTAGAACAGAGGCCCTGGGTGAGCGGCTTTTGTTCAGATGCCGGCCAC | 2831 |
| Qy | 2783 | AACHTCCCTGTAGAACAGAGGCCCTGGGTGAGCGGCTTTTGTTCAGATGCCGGCCAC | 2842 |
| Db | 2832 | GGCTATTTCCTGTTGGTGGCGCTGCTGTGATACCCGGACCCCTGGAGGTCAGAGCGAC | 2891 |
| Qy | 2843 | GGCTATTTCCTGTTGGTGGCGCTGCTGTGATACCCGGACCCCTGGAGGTCAGAGCGAC | 2902 |
| Db | 2892 | TACTCCAGCTATGCCGGACCTCCATCAGAGCAGTCTCACCTTCAACGGCGGCTTCAAG | 2951 |
| Qy | 2903 | TACTCCAGCTATGCCGGACCTCCATCAGAGCAGTCTCACCTTCAACGGCGGCTTCAAG | 2962 |
| Db | 2952 | GCTGGGAGAACATGGGTGCGAAACTCTTTTGGGGTCTTGGGCTGAAGTGTACAGCCGT | 3011 |
| Qy | 2963 | GCTGGGAGAACATGGGTGCGAAACTCTTTTGGGGTCTTGGGCTGAAGTGTACAGCCGT | 3022 |
| Db | 3012 | TTTCTGATTTGCAAGTGACACGCTCCAGACGCTGTGCACCAACATCTACAAGATCCCT | 3071 |
| Qy | 3023 | TTTCTGATTTGCAAGTGACACGCTCCAGACGCTGTGCACCAACATCTACAAGATCCCT | 3082 |
| Db | 3072 | CTGCTGCAGGCGTACAGTTTTCACGCATGTGTCTGACGCTCCCATTTTCATCAGCAAGT | 3131 |
| Qy | 3083 | CTGCTGCAGGCGTACAGTTTTCACGCATGTGTCTGACGCTCCCATTTTCATCAGCAAGT | 3142 |
| Db | 3132 | TGGAAGAACCCCAATTTTTCCTGCGCGTCACTCTGACACGGCTCCCTCTGCTACTCC | 3191 |
| Qy | 3143 | TGGAAGAACCCCAATTTTTCCTGCGCGTCACTCTGACACGGCTCCCTCTGCTACTCC | 3202 |
| Db | 3192 | ATCTCTGAAGCCCAAGAACGAGGATGTGCTGCGGGGCCAAGGGCGCGCGGCCCTCTG | 3251 |
| Qy | 3203 | ATCTCTGAAGCCCAAGAACGAGGATGTGCTGCGGGGCCAAGGGCGCGCGGCCCTCTG | 3262 |
| Db | 3252 | CCCTCCGAGGCGGTGACGTGGCTGTGCCACCAAGCATTCCTGCTCAAGTGTACTCGACAT | 3311 |
| Qy | 3263 | CCCTCCGAGGCGGTGACGTGGCTGTGCCACCAAGCATTCCTGCTCAAGTGTACTCGACAT | 3322 |
| Db | 3312 | CGTGTCACTACGTGCCATCTCTGGGGTCACTCAGGACAGCCAGACGACGTGAGTCCG | 3371 |
| Qy | 3323 | CGTGTCACTACGTGCCATCTCTGGGGTCACTCAGGACAGCCAGACGACGTGAGTCCG | 3382 |
| Db | 3372 | AAGTCCCGGGGAGCGCTGACTGCGCTGGAGCGCGACCAACCCGGCTACTCGCTCA | 3431 |
| Qy | 3383 | AAGTCCCGGGGAGCGCTGACTGCGCTGGAGCGCGACCAACCCGGCTACTCGCTCA | 3442 |

| | | | | |
|----|------|-----|---|------|
| Db | 1701 | GC | CCCCCGCCACACGCTAGTGGACCCCGAAGGCGTCTGGATGCGAAGCGGCGCTGGAACCA | 1760 |
| Qy | 622 | GC | CCCCCGCCACACGCTAGTGGACCCCGAAGGCGTCTGGATGCGAAGCGGCGTGAACCA | 681 |
| Db | 1761 | TAG | CGTCAAGGAGCGCGGGTCCCTCGGGCTGCGACGCCCGGGTGGAGAGGCGCGG | 1820 |
| Qy | 682 | TAG | CGTCAAGGAGCGCGGGTCCCTCGGGCTGCGACGCCCGGGTGGAGAGGCGCGG | 741 |
| Db | 1821 | GG | CAGTGCACGCGAAGTCTGCCGTGCCCAAGAGGCCACAGCGTGGCGTGCCTCTGA | 1880 |
| Qy | 742 | GG | CAGTGCACGCGAAGTCTGCCGTGCCCAAGAGGCCACAGCGTGGCGTGCCTCTGA | 801 |
| Db | 1881 | GC | CGAGCGACGCGCGTTGGGACAGGGTCTCTGGGCCACCCGGGACGAGCGTGACCC | 1940 |
| Qy | 802 | GC | CGAGCGACGCGCGTTGGGACAGGGTCTCTGGGCCACCCGGGACGAGCGTGACCC | 861 |
| Db | 1941 | GAG | TGACCGTGTCTTGTTGTGTGTGTCACCTGCCACAGCCCGCGAAGAGCCACTCTTT | 2000 |
| Qy | 862 | GAG | TGACCGTGTCTTGTTGTGTGTGTCACCTGCCACAGCCCGCGAAGAGCCACTCTTT | 921 |
| Db | 2001 | GG | AGGTCGCTCTCTGGCACGCGCACTCCACCCATCCGTGGGCGGCCACGACACGC | 2060 |
| Qy | 922 | GG | AGGTCGCTCTCTGGCACGCGCACTCCACCCATCCGTGGGCGGCCACGACACGC | 981 |
| Db | 2061 | GG | CCCCCCCCATCATCGGGGCCACCGTCCCTGGGACAGCGCTTGTCGCCGCTGA | 2120 |
| Qy | 982 | GG | CCCCCCCCATCATCGGGGCCACCGTCCCTGGGACAGCGCTTGTCGCCGCTGA | 1041 |
| Db | 2121 | CG | CGAGACCAAGCACTTCTCTACTCTCTACGGCACAGGACGAGCTGGGCGCTCTT | 2180 |
| Qy | 1042 | CG | CGAGACCAAGCACTTCTCTACTCTCTACGGCACAGGACGAGCTGGGCGCTCTT | 1101 |
| Db | 2181 | CC | TACTCAGCTCTCTGAGGCCACGCTGACTGGCGCTCGGAGGTCGTGGAGACCATTT | 2240 |
| Qy | 1102 | CC | TACTCAGCTCTCTGAGGCCACGCTGACTGGCGCTCGGAGGTCGTGGAGACCATTT | 1161 |
| Db | 2241 | TC | TGGGTTCCAGGCGCTGGATGCCAGGACTCCCCGCAAGTTGCCCGCTGCCCCACGC | 2300 |
| Qy | 1162 | TC | TGGGTTCCAGGCGCTGGATGCCAGGACTCCCCGCAAGTTGCCCGCTGCCCCACGC | 1221 |
| Db | 2301 | CT | ACTGCAATCGCGCCCTGTTTCTGAGCTGCTTGGGAACACGCGAGTGGCCCTA | 2360 |
| Qy | 1222 | CT | ACTGCAATCGCGCCCTGTTTCTGAGCTGCTTGGGAACACGCGAGTGGCCCTA | 1381 |
| Db | 2361 | CG | GGTCTCTCTCAAGACGACTGCCCGCTCGAGCTGCGGTCACTCCAGCAGCGGCTG | 2420 |
| Qy | 1382 | CG | GGTCTCTCTCAAGACGACTGCCCGCTCGAGCTGCGGTCACTCCAGCAGCGGCTG | 1341 |
| Db | 2421 | CT | TGCCCGGAGAGCCCGAGGCTCTGTGGCGCCCGCCAGGAGGAGACACAGACC | 2480 |
| Qy | 1342 | CT | TGCCCGGAGAGCCCGAGGCTCTGTGGCGCCCGCCAGGAGGAGACACAGACC | 1401 |
| Db | 2481 | CC | GTGCGCTGGTGCAGCTGCTCGGCCACAGCAGCCCCCTGGCAGGTGTACGGCTTCG | 2540 |
| Qy | 1402 | CC | GTGCGCTGGTGCAGCTGCTCGGCCACAGCAGCCCCCTGGCAGGTGTACGGCTTCG | 1461 |
| Db | 2541 | CG | GGCTGTGCTGCGCGGCTGGTGCCCGCCAGGCTCTGGGGTCCAGGCACACGAACG | 2600 |
| Qy | 1462 | CG | GGCTGTGCTGCGCGGCTGGTGCCCGCCAGGCTCTGGGGTCCAGGCACACGAACG | 1521 |
| Db | 2601 | CC | GTTCTCTCAGGAACACCAAGTTTCATCTCCCTGGGGAGCATGCAAGCTCTCGCT | 2660 |
| Qy | 1522 | CC | GTTCTCTCAGGAACACCAAGTTTCATCTCCCTGGGGAGCATGCAAGCTCTCGCT | 1581 |
| Db | 2661 | GC | AGGACTCAGCTGGAAGATAGCGTGGGGGCTGCCCTTGGCTGGCAGGAGCCGAGG | 2720 |
| Qy | 1582 | GC | AGGACTCAGCTGGAAGATAGCGTGGGGGCTGCCCTTGGCTGGCAGGAGCCGAGG | 1641 |
| Db | 2721 | GG | TGGCTGTGTTCCGGCGCAGACACCGTCTGGTCTGAGGAGATCTGGCAGGTTCT | 2780 |
| Qy | 1642 | GG | TGGCTGTGTTCCGGCGCAGACACCGTCTGGTCTGAGGAGATCTGGCAGGTTCT | 1701 |

| | | | | | |
|----|------|-----------------------|---|-----------------|------|
| Db | 2791 | GCATCGCTGATGANGTGTGTA | CGTCTGAGCTGCTCAGGCTCTTCTCTTTATGTACAGGA | 2841 | |
| Qy | 1702 | GCATCGCTGATGAGTGTGTG | CTGCTGAGCTGCTCAGGCTCTTCTTTATGTACAGGA | 1761 | |
| Db | 2841 | GACCACGTTTCAAAGAACAGG | CTCTTTTCTACCGGAAGAGTGTCTGGACGAAGTTGCA | 2900 | |
| Qy | 1762 | GACCACGTTTCAAAGAACAGG | CTCTTTTCTACCGGAAGAGTGTCTGGACGAAGTTGCA | 1821 | |
| Db | 2901 | AAGCATTTGGAATCAGACAG | CACTTTGAAGAGGTTGACGTCGGGAGCTGTCTCGGAAGCAGA | 2960 | |
| Qy | 1822 | AAGCATTTGGAATCAGACAG | CACTTTGAAGAGGTTGACGTCGGGAGCTGTCTCGGAAGCAGA | 1881 | |
| Db | 2961 | GGTCAGCAGCATCGGAAGCA | CGAGCCGCCCTGTGTAGCTCCAGACTCCGCTTCATCCC | 3020 | |
| Qy | 1882 | GGTCAGCAGCATCGGAAGCA | CGAGCCGCCCTGTGTAGCTCCAGACTCCGCTTCATCCC | 1941 | |
| Db | 3021 | CAAGCCTGACGGCTCGGGCC | GATTCTGGAACATGACACTACGTCGTGGGAGCCAGAACGTT | 3080 | |
| Qy | 1942 | CAAGCCTGACGGCTCGGGCC | GATTCTGGAACATGACACTACGTCGTGGGAGCCAGAACGTT | 2001 | |
| Db | 3081 | CCGCAGAGAAAGAGGCGCG | TCTACCTTCGAGGGTGAAGGCACATGTTTCAGCGTGCT | 3140 | |
| Qy | 2002 | CCGCAGAGAAAGAGGCGCG | TCTACCTTCGAGGGTGAAGGCACATGTTTCAGCGTGCT | 2061 | |
| Db | 3141 | CAACTACAGCGGGCGGGCG | CCCCCGCCCTCTGTGGCGCTCTGTGCTGGGCCCTGGACGA | 3200 | |
| Qy | 2062 | CAACTACAGCGGGCGGGCG | CCCCCGCCCTCTGTGGCGCTCTGTGCTGGGCCCTGGACGA | 2121 | |
| Db | 3201 | TATCCACAGGCGCTGGCGA | CACTTCGTGCTGTGGGGCCCAAGGACCCGCGCCCTGA | 3260 | |
| Qy | 2122 | TATCCACAGGCGCTGGCGA | CACTTCGTGCTGTGGGGCCCAAGGACCCGCGCCCTGA | 2181 | |
| Db | 3261 | GCTGTACTTTGTCAAGTGT | GATGTGACGGCGCTACGACACCATCTCCCGCAGGACAGGCT | 3320 | |
| Qy | 2182 | GCTGTACTTTGTCAAGTGT | GATGTGACGGCGCTACGACACCATCTCCCGCAGGACAGGCT | 2241 | |
| Db | 3321 | CACGGAGGTCAATCCGACGA | TCATCAAAACCCAGAACACGTA | CGTCTGCGTCTCTAC | 3380 |
| Qy | 2242 | CACGGAGGTCAATCCGACGA | TCATCAAAACCCAGAACACGTA | CGTCTGCGTCTCTAC | 2301 |
| Db | 3381 | CGTGGTCCAGAGGCGCCCA | TGGGCGACGTCGCGAAGGCCCTTCAAGAGCCACGCTCTAC | 3440 | |
| Qy | 2302 | CGTGGTCCAGAGGCGCCCA | TGGGCGACGTCGCGAAGGCCCTTCAAGAGCCACGCTCTAC | 2361 | |
| Db | 3441 | CTTGACAGACTCCAGCCGT | TACATGGACAGTTCGTGGCTCACCTGACGAGACAGCAGCC | 3500 | |
| Qy | 2362 | CTTGACAGACTCCAGCCGT | TACATGGACAGTTCGTGGCTCACCTGACGAGACAGCAGCC | 2421 | |
| Db | 3501 | GCTGAGGGATGCCGTGCT | CAATGAGAGAGCTCTCCCTGTAATGAGGCCACAGTGCGCT | 3560 | |
| Qy | 2422 | GCTGAGGGATGCCGTGCT | CAATGAGAGAGCTCTCCCTGTAATGAGGCCACAGTGCGCT | 2481 | |
| Db | 3561 | CTTCGAGCTCTTCTACGCT | TCAATGTCACACCGCGTGGGCATCAGGGCGAAGTCTCTA | 3620 | |
| Qy | 2482 | CTTCGAGCTCTTCTACGCT | TCAATGTCACACCGCGTGGGCATCAGGGCGAAGTCTCTA | 2541 | |
| Db | 3621 | CGTCCAGTGGCAGGGATCC | CGCAGGGCTCAATCTCTCCACGCTGTCTGACGCGTGTG | 3680 | |
| Qy | 2542 | CGTCCAGTGGCAGGGATCC | CGCAGGGCTCAATCTCTCCACGCTGTCTGACGCGTGTG | 2601 | |
| Db | 3681 | CTACGGCGACATGGAAACA | AGCTGTTTTCGGGGGATTCGGGGGACGGGCTGCTCTCGG | 3740 | |
| Qy | 2602 | CTACGGCGACATGGAAACA | AGCTGTTTTCGGGGGATTCGGGGGACGGGCTGCTCTCGG | 2661 | |
| Db | 3741 | TTTGGTGGATGATTTCTGT | TGTGTGACACCTCACTCACCACGGGAAACCTTCCTCTCAG | 3800 | |
| Qy | 2662 | TTTGGTGGATGATTTCTGT | TGTGTGTGTGACACTCACCTCACCCACGCGAAACCTTCCTCTCAG | 2721 | |
| Db | 3801 | GACCCCTGTCGAGGCTGCC | TGATGGCTGCTGTGACTTGTGCGAAGACAGTGGT | 3860 | |
| Qy | 2722 | GACCCCTGTCGAGGCTGCC | TGATGGCTGCTGTGACTTGTGCGAAGACAGTGGT | 2781 | |
| Db | 3861 | GAACCTCCCTGTGAGAAGA | CAGAGCCCTGGGTGGCACGGCTTTTGTTCAGATGCCGCGCCA | 3920 | |

```
Qy 2782 GAACCTCCCTGTAGAACGAGGAGCCCTGGGTGGACAGCGCTTTGTTGATGCGCGGCCA 2841
Db 3921 CGGCCTATTCCTCCCTGGTGGCGCTGCTGCTGGATACCCGCGACCTGAGGTGCGAGCGGA 3980
Qy 2842 CGGCCTATTCCTCCCTGGTGGCGCTGCTGCTGGATACCCGCGACCTGAGGTGCGAGCGGA 2901
Db 3981 CTACTCCAGCTATGCCCCGAGACTCCATCAGAGCCAGTCTCAGCTTCAACCGCGGCTTCAA 4040
Qy 2902 CTACTCCAGCTATGCCCCGAGACTCCATCAGAGCCAGTCTCAGCTTCAACCGCGGCTTCAA 2961
Db 4041 GGTGGGAGGAGACATGCGTCCGCAACTCTTTGGGCTCTTGGGCTTGGGCTGAGTGTCCAGCGCT 4100
Qy 2962 GGTGGGAGGAGACATGCGTCCGCAACTCTTTGGGCTCTTGGGCTTGGGCTGAGTGTCCAGCGCT 3021
Db 4101 GTTCTGGATTGTCAGGTGAACAGCTCCAGAGCGTGTGCGACCAACATCTACAAGATCCT 4160
Qy 3022 GTTCTGGATTGTCAGGTGAACAGCTCCAGAGCGTGTGCGACCAACATCTACAAGATCCT 3081
Db 4161 CTGCTCAGGGGTACAGGTTTACAGGATGTGCTGAGCTCCATTTTCATCAGCAAGT 4220
Qy 3082 CTGCTCAGGGGTACAGGTTTACAGGATGTGCTGAGCTCCATTTTCATCAGCAAGT 3141
Db 4221 TTGGAAGAACCCACATTTTCTCGCGTCTATCTCTGACAGCGCTCCCTCTGCTACTC 4280
Qy 3142 TTGGAAGAACCCACATTTTCTCGCGTCTATCTCTGACAGCGCTCCCTCTGCTACTC 3201
Db 4281 CATCTCTGAAGCAAGAACGAGGATGTGCTGGGGGCCAAGGGCGCGCGGCCCTCT 4340
Qy 3202 CATCTCTGAAGCAAGAACGAGGATGTGCTGGGGGCCAAGGGCGCGCGGCCCTCT 3261
Db 4341 GCCCTCCAGGGCGTACAGTGTGCTGCCACCAAGCATTTCTGCTCAAGCTGACTCGACA 4400
Qy 3262 GCCCTCCAGGGCGTACAGTGTGCTGCCACCAAGCATTTCTGCTCAAGCTGACTCGACA 3321
Db 4401 CGGTGTACCTACGTGCGACCTCTCGGGTCACTAGGACAGCCAGAGCGAGCTGAGTCG 4460
Qy 3322 CGGTGTACCTACGTGCGACCTCTCGGGTCACTAGGACAGCCAGAGCGAGCTGAGTCG 3381
Db 4461 GAAGCTCCCGGGGAGCGTACTGCTGAGCGCCGAGCGCCAGCCACCGGCGACTGCCCTC 4520
Qy 3382 GAAGCTCCCGGGGAGCGTACTGCTGAGCGCCGAGCGCCAGCCACCGGCGACTGCCCTC 3441
Db 4521 AGACTTCAAGACCATCTCTGGACT 4543
Qy 3442 AGACTTCAAGACCATCTCTGGACT 3464

RESULT 4
LOCUS AF073311 3369 bp mRNA ROD 09-SEP-1998
DEFINITION Mus musculus telomerase catalytic subunit mRNA, complete cds.
ACCESSION AF073311
NID g3551846
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 3369)
TITLE Expression of mouse telomerase catalytic subunit in embryos and
adult tissues
JOURNAL Martin-Rivera, L., Herrera, E., Albar, J. P. and Blasco, M. A.
MEDLINE 98393668
REFERENCE 2 (bases 1 to 3369)
AUTHORS Martin-Rivera, L., Herrera, E. and Blasco, M. A.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-1998) Immunology and Oncology, National Centre of
Biotechnology, Cantoblanco, Madrid 28049, Spain
FEATURES
1. 3369
source /organism="Mus musculus"
/db_xref="taxon:10090"
```

```
CDS
1..3369
/codon_start=1
/product="telomerase catalytic subunit"
/db_xref="PID:g3551847"
/translation="MTRAPRCAPVRSLLRSRVVWPLATFVRRLGEGRRLLVOPGDP
KLYRTVAOCLVCMHMGSPPPADLSFHOVSILKELVAVRVORLCERNERNVLAQFE
LNEARGGPPMAFTSSVRSYLFNTVETLRVSGANWMLLSRVDDLLYLLAHALYL
LVPPSCAYQVCGSPLYICATDDIMPVSASIRPTRPVGRNFTNLRLFLQOIKSSRQE
AKPLALPSRGTKRHLSLSTSVPSAKKARCVPRVEEGPHRQVLPTSGKSWPSP
ARSPVPTAEKDLSSKGVSDLSGSCVCKKHPSTSLSPRQNAFOLRPIETFRH
FLYSRDGOERLNPSFLLNLOPLNTGARRVEIIFLGSRPRTSGPLCTRHLRSRPMQ
OMRPLFOQLLVNHAECOVYRLRSHCRFTANOQVTDALNTPPHLMDLLHLSSPMQ
VYGFRLACLCKVVSASLWGTNRNRRFFNKKFISLGKIGLSQLSGLMWMKRYEBC
WURSSGKDRVPAARHRLERILATFLWMDTYVQLLRSFFYITESTFQNRLEFY
KUSSWSKLSIGVROHLRLVRELSQEERHHQDTWLAMPICRLRIFPKNGRLIV
NYSYSGTGRALGRKQAOHTQRLKTLFSLNAYERTKHPHMGSSVLGNDIYRTWRA
FYLVRALDQTPMYEVKADVTGAYDAIPOGKLEVVANNIRHSESTYCIROYAVVR
DSOGVHKFSRQVTTLSLQPYMGOFKHLQDSASALRNSVVIQSIISMNESSSL
FDFFLRHSHVVKIGRCYTCQGPQSGSSLSLCSLCFGDMENKLEAEVQORDQL
LRVDDFLVTPHLDQAKTFLSLVHGVPEYGCMLNKTVMNFPVPEPTGLGAAPIY
LPAHCLFPWCGLLDQTLLEVCFDYSYQATSIKTSLTFSQVFKAGTKMRNKLVLIR
LKCHGLFLQVNSLTQVCINIKIFLQAYRFHACVIOLPDQVRKMLTFLGIIS
SOASCYAILKVKNPMTLKASGPPEAAHWLCYQAFLLKLAHSHVYIKCLLGPLRT
AOKLCLCKLPEATMTILKAAADPALSTDFTILD"
BASE COUNT 737 a 959 c 891 g 782 t
ORIGIN
Query Match 28.4%; Score 1079; DB 28; Length 3369;
Best Local Similarity 68.2%; Pred. No. 0.00e+00;
Matches 2343; Conservative 0; Mismatches 988; Indels 102; Gaps 6;
Db 1 ATGACCGCGCTCTCTGCTGCCCGCGGTGGCTCTCTGTCGCGACCGCATCCGGAG 60
Qy 68 ATGCGCGCGCTCCCCGCTGCGAGCGGTGGCTCTCTGTCGCGACCGCATCCGGAG 127
Db 61 GTGTGGCGCTGCAACCTTTGTCGCGCGCTGCGGCGCGCGCGCGCGCTGTGTCAA 120
Qy 128 GTGTGGCGCTGCGCGCGCTGCGCGCGCTGCGGCGCGCGCGCGCGCTGCGCGCT 187
Db 121 CCGCGGACCGAAGATCTACCGCCTTTGGTGGCCCAATGCCTAGTGTGCATGCACTG 180
Qy 188 CGCGGGACCGCGCGGCTTTCCGCGCGCTGTCGCGCGCGCTGTCGCTGCGCTGG 247
Db 181 GGCTACAGCGCTCCACCTGCGGACCTTTTCCTTCACAGGTGTGTCATCCCTGAAAGACTG 240
Qy 248 GACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 307
Db 241 GTGGCGAGGTTGTGACAGACTCTGCGAGCGCAACGAGACGAGACGCTGGCTTTGGC 300
Qy 308 GTGGCGCGAGTGTGACAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCT 367
Db 301 TTTGAGCTGCTTAACGAGCGAGCGCGCGCTCCCATGGCCTTCACTAGTAGCGTCCGT 360
Qy 368 TTCGCGTGTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGC 427
Db 361 AGCTACTTGGCCCAACACTGTTATTAGACCGCTGCGTGTGTCAGTGTGTCATGATGCTACTG 420
Qy 428 AGCTACTTGGCCCAACCGGTGACCGAGCTGCGGGGAGCGGGGCGGTGCGGGGCTGCTG 487
Db 421 TTGAGCGAGTGGCGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 488 CTGCGCGCGTGGCGCGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 547
Db 481 CTGCTGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 548 CTGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 607
Db 541 ACCACGATATCTGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy 608 GCCACTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 665
Db 601 AATTTCACCTTAGTCTTACACAGATCAAGACGAGTAGTCGCCAGGACGACCG 660
```

666 -----AACGGCCTGGAACCATAGCGTCAGGGCGCGG 700
Db 661 AAACCCCTGGCCTTGCCATCTCGAGGTACAAGAGGCACTGAGTCTACACAGTACAAGT 720
QY 701 GTCCCCCTGGCCTTGCCAGCCCGGGTGCAGGAGGCGCGGGGCGAGTGCACCGCAAGT 760
Db 721 GTGCTTACAGTAAGAAGGCCAGATGCTATCTGTCTCCGAGAGTGGAGGAGGCCGCCAC 780
QY 761 CTGCGGTGCCCAGAGGCCAGCGTGGCGCTGCCCTGAGCCGGAGCGGAGCGCCGTT 820
Db 781 AGGAGGTGCTACCAACCCCATCAGGCAAAATCATGGTGCCAAAGTCTCTGCTCGGTCCCC 840
QY 821 GGGCAGGGGTCTCTGGGCCCAACCGGCGAGGACGCGTGGACCGAGTACCGTGTCTGT 880
Db 841 GAGGTGCTTACTCAGAGAAAGATTTGTCTTCTTAAGGAAGGTGTGTGACCTGAGTCTC 900
QY 881 GTGGTGTCACTGCCAGACCGCGGGAAGAACCACTCTTTGGAGGGTGGCCTCTCTGGC 940
Db 901 TCTGGG-----TCGGTGTGCTGTAACACAAAGCCAGCTCCACATCTCTGCTG 948
QY 941 ACGGGCCACTCCCAACCATCCGTGGGCGCGCAGCACACCGGGCCCCCATCCACATCG 1000
Db 949 TCACCCCGCCGCAAAATGCTTACGCTCAGGCCAATTTATTGAGACCAAGACATTTCCCT 1008
QY 1001 CGGCCACCACTCCCTGGGACAGCCTTGTCCCGGTGTACGCCGAGACCAAGCACTTC 1060
Db 1009 TACTCAGGGGAGATGCCAAGAGCGTCTAAACCCCTCATCTTCTACTCAGACACCTCCAG 1068
QY 1061 CTCTACTCTCCACGCGCAAGAGGACAGCTCGCGCCCTCTCTTCTACTCAGCTCTCTGAG 1120
Db 1069 CTTAACTTGACTGGGGCAGGAGCTGCTGAGATCATCTTTCTGGCTCAAGGCCCTAGG 1128
QY 1121 CCAGCCTGACTCGGCTCGGAGGCTCGTGAGACCACTTTCTGGGTTCAGGCCCCGTG 1180
Db 1129 ACATCAGGACCACTCTCAGGACACACCGCTCTATCGGCTCGATCTGGCAGATCGCGCCC 1188
QY 1181 ATGCCAGGACTCCCCCAGCTTGGCCCGCTTCCCCAGCGCTACTGGCAATGGCGCC 1240
Db 1189 CTGTTCCACAGCTGCTGTGTAACCATGACAGTGCACAAATATGTACAGTCTCTCAGGTCA 1248
QY 1241 CTGTTCTGGAGTGTCTGGGAACACAGCGCAGTGCCTTACGGGGTGTCTCTCAAGACG 1300
Db 1249 CATTGAGGTTTCGACAGCAACCAACAGTGCAGATGCCCT----- 1292
QY 1301 CACTGCCGCTGCGAGTGGGTACCCCGCAGCGCGGTGTCTGTGCCCGGGAGAGCC 1360
Db 1293 -----GAACACCAAGCCCGCAGCCACCGCCTCATGGATTG 1323
QY 1361 CAGGGCTCTGTGGGCCCCGAGGAGGAGACACAGACCCCGTCCCTTGGTGCACTG 1420
Db 1324 CTCGGCTGCACAGCAGTCCCTGGCAGGTATATGGTTTTCTTCGGGCGCTCTCTGCAAG 1383
QY 1421 CTCGGCAGCACAGCAGCCCTGGCAGGTGTACGGCTTCGTGGGGCTGCCTGGCGCG 1480
Db 1384 GTGTTGCTCTAGTCTCTGSGGTACAGGCACAATGAGCGCGCTCTTTAAGACTTA 1443
QY 1481 CTGGTGCCCCAGGCTCTGGGCTCCAGGCACAAGACCGCGCTTCTCAGGAACACC 1540
Db 1444 AAGAAGTTTCACTCTGTTGGGAAATACGCCAAGCTATCACTGAGGAACCTGATGTGGAAG 1503
QY 1541 AAGAAGTTTCACTCTGTTGGGAGCATGCCAAGCTCTCGTGCAGGAGCTGACGTGGAAG 1600
Db 1504 ATGAAGTAGAGGATTGCCACTGGCTCCGACGAGCCCGGGGAGGACCGTGTCCCCGCT 1563
QY 1601 ATGAGCGTGGGACTCGCTTGGCTCGCAGGAGCCAGGGTGTGGCTGTGTTCGGCC 1660
Db 1564 GCAGACACCGCTGAGGGAGAGATCTGCTACGTTCTCTGCTGTCTGCGCTGAGACACA 1623
QY 1661 GCAGAGCACCGCTGCGGTGAGGAGATCCTGGCCAAGTCTGTGCACTGGCTGATGAGTGTG 1720
Db 1624 TAGCTGTACAGTGTCTAGGTCTATTTTATCATCAGACAGACACATTCAGAGAGAAC 1683
QY 1721 TAGCTGTACAGTGTCTAGGTCTTTCTTTTATGTACGGAGACCAAGTTTCAAAAGAAC 1780

Db 1684 AGGCTCTTCTTACCGTAAGAGTGTGTGGAGCAAGCTGCAGAGCATTTGGAGTCAAGCAA 1743
QY 1781 AGGCTCTTCTTACCGTAAGAGTGTGTGGAGCAAGTGTGCAAGCATTTGCAATCAGACAG 1840
Db 1744 CACTTTGAGAGAGTGTGGGTACGGGAGCTGTCAACAAGAGAGGTCAAGCATACCAAGGAC 1803
QY 1841 CACTTGAAGAGGTGTGAGTTCGGGAGCTGTCTGGAAGAGAGGTCAAGCAGCATCGGGAA 1900
Db 1804 ACCTGCTAGCCATGCCCATCTGCAGACTGCGCTTATCCCCAAGCCCAACAGCCCTGCGG 1863
QY 1901 GCCAGCCCGCCCTGTGAGTCCAGACTCCGCTTATCCCCAAGCTGACGGGCTGCGG 1960
Db 1864 CCATTGTGAACATGAGTTATAGCATGGGTACCAGAGCTTTTGGCGAGAGAGACAGGCC 1923
QY 1961 CCGATTGTGAACATGAGTACGTCGTGGGAGCCAGAACGTTCCGACAGAAAAGAGGCC 2020
Db 1924 CAGCATTTACCCAGCGCTCTCAAGACTCTCTTCAGCATGCTCAACTATGACCGGCAAAA 1983
QY 2021 GAGCGTCTCACTCGAGGTGAAGGACTGTTCAAGGTGCTCACTACGACGCGGCGCG 2080
Db 1984 CATCTCACTTATGGGGTCTTCTGACTGGGTATGAATGACATCTACAGAGCTGGCGG 2043
QY 2081 CGCCCGGCTCTCGGCGCTCTGTGCTGGGCTTGACGATATCCACAGGCGCTGGCGC 2140
Db 2044 GCCTTGTGCTGCTGTGCGTCTGAGACACACACAGGATGACTTTGTTAAGGCA 2103
QY 2141 ACCTTGTGCTGCTGTGGGGCCAGGACCCCGCCTGAGCTGACTTTGTCAAGGTG 2200
Db 2104 GATGTACCGGGGCTTATGATGCCATCCCGCAGGTAAGCTGGTGAGGTGTGTGCCAAT 2163
QY 2201 GATGTACGGGCGCTGACGACACCATCCCCCAGGAGCGCTCACGGAGTCTATCGCCAGC 2260
Db 2164 ATGATCAGGCACTCGGAGAGCAGCTACTGTATCCGCGAGTATGCACTGGTGGCGAGAT 2223
QY 2261 ATCATCAACCCC--AGAACAGTACTGCGTGGGTATGCCGTGCTTCAGAGGCC 2317
Db 2224 AGCCAAGGCCAAGTCCACAGTCTCTTTAGAGACAGGTACACACCCTCTCTGACCTCCAG 2283
QY 2318 GCCATGGGCACGTCGCCAAGGCTTCAAGAGCCAGCTCTTACCTTGACAGACCTCCAG 2377
Db 2284 CCATACATGGCCAGTCTCTTAAGCATCTGCAGGATTCAGATGCCAGTGCACGTGAGGAAC 2343
QY 2378 CCGTATACGGAGACTGCTGTGCTCACCTGCAGGAGAC-----CAGCCGCTGAGGGAT 2431
Db 2344 TCCGTTGTCTATCAGCAGAGCATCTCTATGAATGAGAGCAGCAGCAGCGCTGTTTGAATTC 2403
QY 2432 GCGTGTCTATCGAGCAGAGCTCTCTCCCTGAATGAGCCAGCAGTGGCTCTTCGAGCTC 2491
Db 2404 TCTCTGCATCTCTGCTCACAGTGTCTGTAAGATTGGTGACAGGTGCTATACGCAAGTC 2463
QY 2492 TCTCTACGCTTCTATGTGCCACACCGCGTGGCATCAGGGGCAAGTCTTACGTCCAGTGC 2551
Db 2464 CAGGGATCCCCCAGGCTCCAGCCTATCCACCCTGCTGTGAGTCTGTGTTTCGGAGAC 2523
QY 2552 CAGGGATCCCGCAGGGCTCCATCTCTCCACGCTCTCTGACGCTGTGCTACGCGCAG 2611
Db 2524 ATCGAACAACAGTGTGCTGAGGTGCAGCGGATGGGTGCTTTTACGTTTTGTGTAT 2583
QY 2612 ATCGAACAACAGTGTGTTGGGGGATTCGGGGGAGCGGGTGTCTGCTGTTGTGGAT 2671
Db 2584 GACTTTCTTGTGTGAGCCTCACTTTGGACCAAGCAAAAACCTTCTCAGCACCCCTGGTC 2643
QY 2672 GATTTCTTGTGTGACACTCACTCACTCACCCACGCGAAAACCTTCTCAGGACCCCTGGTC 2731
Db 2644 CATGGGTTCTCTGAGTATGGGTGCATGATAACTTCGAGAAGACAGTGGTGAATCCCT 2703
QY 2732 CGAGGTCTCTCTGAGTATGGCTGCTGGTGAAGTTCGGAAGACAGTGGTGAATCCCT 2791
Db 2704 GTGAGCCTGGTACCCTGGGTGGTGCAGTCCATACACAGTGCCTGCTCAGTGCCTGTT 2763
QY 2792 GTAGAAGACAGGCCCTGGGTGGCAGCGCTTTTGTTCAGATGCCGCCACCGCCCTATTTC 2851

| | | | |
|--|------|--|------|
| Db | 2764 | CCCTGGTGGGCTTGTCTGTGGACACTCAGACTTTGGAGAGTGTCTGTGACTACTCAGGT | 2823 |
| QY | 2852 | CCCTGGTGGGCGTGTCTGTGGATACCCGGACCGTGGAGTGCAGAGCGACTACTCCAGC | 2911 |
| Db | 2824 | TATGCCACAGACCTCAATTAGAGGAGCCTCAGCTTCCAGAGTGTCTTCAAGAGCTGGGAAG | 2883 |
| QY | 2912 | TATGCCGGGACCTCCATCAGACGAGTCTCACTTCAACCGGGCTTCAAGGCTGGGAGG | 2971 |
| Db | 2884 | ACCATCGGGAACAAGCTCCTGTCTGGTCTTGGCGTTGAAGTGTACGGTCTATTTCTAGAC | 2943 |
| QY | 2972 | AACATGGCTGCAAACTCTTTGGGTCTTGGCGCTGAAGTGTACAGCGCTGTTCTGGAT | 3031 |
| Db | 2944 | TTGCAGGTGAACAGCCTCCAGACAGTCTGCATCAATATATACAAGATCTTCTCTGTTTCAAG | 3003 |
| QY | 3032 | TTGCAGGTGAACAGCCTCCAGAGCGGTGTGCACCAACATCTACAAGATCCTCTGCTGCAG | 3091 |
| Db | 3004 | GCCTACAGGTTCCATGCATGTGTGATTACGCTTCCCTTTGACCAGCGTGTAGGAAGAAC | 3063 |
| QY | 3092 | GCCTACAGGTTTCAACGATGTGTCTGTCAGCTCCCATTTTCATCAGCAAGTTTGAAGAAC | 3151 |
| Db | 3064 | CTCACATTTCTTGGGCACTATCTCCAGCAGCATCTCTGCTATGCTATGCTATCTCTGAAG | 3123 |
| QY | 3152 | CCACATTTTCTGCGCGTCACTCTGACAGCGGCTCCTCTGCTATCTCATCTCTGAAA | 3211 |
| Db | 3124 | GTCAAGAATCCAGGAATGACACTAAAGGCTCTGGCTCTCT-----TTCTCTCTGAA | 3174 |
| QY | 3212 | GCCAAAGAACGAGGATGTCTGTGGGGCCAAAGGCGCGCGGCCCTCTGCCCTCCGAG | 3271 |
| Db | 3175 | GCCGCACATTTGGCTCTGCTACCAAGCTTCTCTGCTCAAGCTGGCTGCTATCTGTGATC | 3234 |
| QY | 3272 | GCCGTGCAGTGGCTGTGCCACCAAGCATTTCTGCTCAAGCTGACTCGACACCGTGTACC | 3331 |
| Db | 3235 | TACAAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACTCTGCTGCCGGAAGCTCCCA | 3294 |
| QY | 3332 | TAGTGCCACTCTGGGGTCACTCAGGACAGCCCAAGCAGCAGCTGAGTCGGAAGCTCCG | 3391 |
| Db | 3295 | GAGGCGCAATGACCATCTTTAAAGTGCAGCTGACCCAGCCCTAAGCACAGACTTTTCAG | 3354 |
| QY | 3392 | GGGACGAGCTGACTGCCCTGGAGGCGGACCAACCGGCACTGCCCTCAGACTTCAAG | 3451 |
| Db | 3355 | ACCATTTTGGACT 3367 | |
| QY | 3452 | ACCATCTGGACT 3464 | |
| RESULT 5 | | | |
| LOCUS AF051911 3426 bp mRNA ROD 02-APR-1998 | | | |
| DEFINITION Mus musculus telomerase reverse transcriptase mRNA, complete cds. | | | |
| ACCESSION AF051911 | | | |
| NID g3005591 | | | |
| KEYWORDS | | | |
| SOURCE house mouse. | | | |
| ORGANISM Mus musculus | | | |
| REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; | | | |
| AUTHORS Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| TITLE 1 (bases 1 to 3426) | | | |
| JOURNAL Greenberg,R.A., Allsopp,R.C., Chin,L., Morin,G.B. and DePinho,R.A. | | | |
| REFERENCE Expression of mouse telomerase reverse transcriptase during | | | |
| AUTHORS Oncogene (1998) In press | | | |
| TITLE 2 (bases 1 to 3426) | | | |
| JOURNAL Greenberg,R.A., Allsopp,R.C., Chin,L., Morin,G.B. and DePinho,R.A. | | | |
| REFERENCE Submitted (02-MAR-1998) Microbiology and Immunology, Albert | | | |
| AUTHORS Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY | | | |
| TITLE 10461, USA | | | |
| FEATURES | | | |
| source Location/Qualifiers | | | |
| 1..3426 | | | |
| /organism="Mus musculus" | | | |
| /db_xref="taxon:10090" | | | |
| /chromosome="13" | | | |
| 30..3398 | | | |
| CDS | | | |
| /map="MGD offset position 40; 40 cm from centromere" | | | |

QY 652 GCGTCTGGATGCG-----AACGGGCTTGAACCATAG 684
Db 674 TCGCAGGAAGCAGCAACCCCTGGCCCTGCCATCTCAGAGGTACAAAGAGGCATCTGAG 733
QY 685 CGTCAGGAGAGCCCGGGTCCCTCGGCTCCCAAGCCCCGGTGCAGAGAGCGCGGGG 744
Db 734 TCTCACCAGTACAAGTGTGCGCTTACGCTTAAGAAGGCCAGATGCTATCTGCTCCCGAGAGT 793
QY 745 CAGTGCCAGCGAAGTCTGCCGTGTGCCCAAGAGGCCAGGCGTGGCGGTGCCCTGAGCC 804
Db 794 GGAGGAGGACCCACAGCAGAGGTGCTTACCACCCCATCAGGCAAAATCATGGGTGCCAAG 853
QY 805 GGAGCGGACGCCGTGGCGCAGGGGTCTCTGGCCCCACCCCGCAGGACGCGTGGACCGAG 864
Db 854 TCCTGTCTCGGTCCCGCAGAGTGCCTACTGTCAGAGAAAGATTGTCTTAAAGAAAGGT 913
QY 865 TGACCGTGGTTCTGTGTGGTGTACCTTGGCCAGACCCCGCAAGAACCCCTCTTTTGA 924
Db 914 GTCTGACCTGAGTCTCTCTGGG-----TCGGTGTGCTGTAAACACAAAGCCAG 961
QY 925 GGGTGGCTCTCTGGCAGCGGCCACTCCACCCATCCGCTGGCGGCCAGCACAGCGGG 984
Db 962 CTCACATCTCTGTGTCACCAACCCGCCAAATGCTTTTTCAGCTCAGGCCATTTATGA 1021
QY 985 CCCCCATCCACATCGCGGCCACACAGCTCCCTGGGACAGCGCTTGTCCCCGGGTAGCG 1044
Db 1022 GACACAGATTTCTTTACTCAGGGGAGATGGCCAGAGCGGTCTAAACCCCTCATTTCT 1081
QY 1045 CGAGACCAAGCACTTCTCTACTCTCAGGGCAAGAGGAGAGTGTGCGGCCCTCTCTCT 1104
Db 1082 ACTCAGCAACCTCAGCGCTAACTTACTGAGTGGGCCAGGAGACTGTGGAGATCATCTTCT 1141
QY 1105 ACTCAGCTCTCTGAGGCCACGCTGACTTGGCGCTCGGAGGCTGTGGAGACCATCTTCT 1164
Db 1142 GGGCTCAAGGCTTAGGACATCAGGACCACTCTGAGGAGCACACCGTCTATCGCGTCGATA 1201
QY 1165 GGGTTCAGGCCCTGGATGCCAGGACTCCCGCAGAGTTGCCCGCTGCCCGAGCGCTA 1224
Db 1202 CTGGCAGATGGGCCCTTGTTCACAGACTGCTGTGTGAACCATGACAGAGTGCCTATGCT 1261
QY 1225 CTGGCAATGGGGCCCTGTTTCTGGAGCTGCTTGGGAACACGCGCAGTGCCTCTACGG 1284
Db 1262 CAGACTCCTCAGGTACATTCAGGTTTTCAGACAGCAACCAAGTGTACAGATGCTT 1321
QY 1285 GGTGCTCTCAAGAGCACTGCCCGCTCGAGCTGCGGTACCCAGCAGCGGTGTGTG 1344
Db 1322 -----GAACACAGCCCAACC 1336
QY 1345 TGCCCGGAGAACCCAGGGCTCTGTGGCGCCCCCAGGAGGAGACAGACCCCGC 1404
Db 1337 GCACCTCAGGATTTGCTCCGCCCTGCACAGCAGTCCCTGGCAGGTATATGTTTCTTCG 1396
QY 1405 TCGCTGTGAGCTGCTTCCGCCAGCAGCAGAGCCCTTGGCAGGTGTACGGTTCGTGCG 1464
Db 1397 GGCCTGTCTGCAAGGTGGTGTCTGTAGTCTCTGGGGTACCAGGCACATGAGCGCG 1456
QY 1465 GGCCTGTCTGCGCGGTGGTGGCCCCAGGCTCTGGGGCTCCAGGCACACAGCAACGCG 1524
Db 1457 CTCTCTTTAAGAACTTAAAGAAAGTTCATCTGTTGGGAAATACGGCAAGCTATCACTCA 1516
QY 1525 CTCTCAGGAACACCAAGATTCATCTCCTGGGGAAGCATGCCAAGCTCTGCTGCA 1584
Db 1517 GGAACCTGATGTGAAGATGAAGTAGAGGATTTGCCACTGGCTCCGACAGCCCGGGAA 1576
QY 1585 GGAGCTGACGTGGAAGATGAGCGTGGCGGACTTGGCTTGGCTGCGAGGACCCAGGGGT 1644
Db 1577 GGACCGTGTCCCGCTGACAGACCCGCTCAGGGAGAGATCTGGGTACGTTCCTGTT 1636
QY 1645 TGGCTGTCTTCGGCGCCGAGAGACCCGCTCTCGTGAGAGATCTTGGCCAAGTTCCTGCA 1704
Db 1637 CTGGCTGATGACACATACGTGGTACAGCTGCTTAGGTCTTCTTTTACACACAGAGAG 1696
QY 1705 CTGGCTGATGAGTGTGTACGTCTGTCGAGCTGCTCAGGTCTTCTTTTATGTCAGGAGAC 1764

Db 1697 CACATTCAGAGAAACAGGCTCTTCTTACCCTAAGAGTGTGTGGAGCAAGCTGCAGAG 1756
QY 1765 CACGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGAGCAAGTTGCAAG 1824
Db 1757 CATTGGAGTCAGGCAACACACCTTGAAGAGTGCCTACGGGAGCTGTACACAAAGAGGAGT 1816
QY 1825 CATTGGATCAGACAGCACTTGAAGAGGTGCAGCTCGGGAGCTGTCCGAGAGAGGT 1884
Db 1817 CAGCATCACAGGACACCTGGCTAGCCATGCCCATCTGCAGACTGGGCTTTCATCCCAA 1876
QY 1885 CAGCAGCATCGGGAAGCCAGGCCCGCTGTGACGTCCAGACTCCGCTTTCATCCCAA 1944
Db 1877 GCCCAAGGGCTCGGGGCCCATTTGTAACATGAGTTATAGCATGGGTACCAGAGCTTTGGG 1936
QY 1945 GCCTGACGGCTCGGGCGGATTTGTAACATGAGTGTGCTGTGGGAGCCAGAGCTTCCG 2004
Db 1937 CAGAAGAACAGGCGCCAGCATTTACCCAGCGCTCTCAAGACTCTCTTACAGCATGCTCAA 1996
QY 2005 CAGAGAAAGAGGCCGCGCTCTACCTCGAGGCTGAGGCATGTTTACAGCTGCTCAA 2064
Db 1997 CTATGAGCGGACAAAACATCCTCACCTTATGGGTCTTCTGTACTGGGTATGAATGAT 2056
QY 2065 CTACAGCGGGCGCGGCCCTTCTGGCGCTCTGTGCTGGGCTGGAGGAT 2124
Db 2057 CTACAGACCTGGCGGGCTTTGTGCTGTGCTGTGCTGTGACACAGACACCCAGAT 2116
QY 2125 CCACAGGCGCTGGCGCACCTTGTGCTGTGCTGTGCGGGCCAGGACCCGCGCTGAGCT 2184
Db 2117 GTACTTTTAAAGCAGATGTACCGGGCTATGATGCCATCCATCCCGAGGTAGCTGGT 2176
QY 2185 GTACTTTGTCAAGTGTGATGTGACGGCGGTACGACACCATCCCGAGGACAGGCTAC 2244
Db 2177 GGAGTGTGTGCCAATATGATCAGGCACTCGGAGAGCACGTACTGTATCCCGCAGTAGC 2236
QY 2245 GGAGTTCATGCCAGCATCATCAAAACCC---AGAACACGTACTGCTGCGTATGCG 2301
Db 2237 AGTGGTCGGAGATAGCCNAGGCCAAGTCCCAAGTCTTTTAGGAGACAGGTACACC 2296
QY 2302 CGTGGTCAGAAGCCCGCCATGGGCACGTCGCGCAAGGCTTCAAGAGCCACGCTCTAC 2361
Db 2297 CCTCTGTACTCCAGCATACATGGCCAGTTCCTTAAAGCATCTGCAGGATTCAGATGC 2356
QY 2362 CTTGACAGCTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTGCGAGAGAC----- 2415
Db 2357 CAGTGCACTGAGGAACCTCCGTTGTCATCGACAGAGATCTCTATGAATGAGACAGCAG 2416
QY 2416 CAGCCCGCTGAGGATGCCGTGCTCATCGAGCAGAGCTCCTCCTGAATGAGGCCAGCAG 2475
Db 2417 CAGCCTGTTGACTTCTCTGCACTTCTGCTGCTCAGAGTGTCTGAAGATTTGTTGACAG 2476
QY 2476 TGGCCTCTTCGACGCTCTTCTACGCTTCATGTGCCACACCCCGTGGCATCAGGGGCA 2535
Db 2477 GTGCTATACAGTGCAGGCGCATCCCCAGGGCTCCAGGCTATCCACCCCTGCTCTGAG 2536
QY 2536 GTCTACGTCCAGTGCAGGGGATCCCCGAGGGGTCCATCTCTCCACGCTCTGCGAC 2595
Db 2537 TCTGTGTTCCGACACATGGAAACAAGCTGTTTGTGAGTGCAGCGGGATGGTGTGCT 2596
QY 2596 CCTGTGCTACGGCACATGAGAACAGAGTGTTCGGGGATTCGGCGGGACGGCTGCT 2655
Db 2597 TTTACGTTTTGTGATGACTTCTGTTGGTACGCCCTCAGTTGGACCAAGCAAAACCTT 2656
QY 2656 CCGTGGTTTGTGTGATGATTTCTTGTGTTGGTGACACCTCACCTCACCACCGGAAACCTT 2715
Db 2657 CCTCAGCACCTGTGCTCATGGCTTCTGAGTATGGGTGATGATAAATCTTGCAGAGAC 2716
QY 2716 CCTCAGACCCCTGTGCTCCGAGGTGCTCCGTGAGTATGGCTGGTGTGCTGCGAAGAC 2775
Db 2717 AGTGGTGAACCTCCTGTGGAGCCTGTGTACCTGGGTGGTGCAGCTCCATACCAAGTGC 2776
QY 2776 AGTGGTGAACCTCCTGTAGNAGAGAGGCGCTGTGGTGGCACGGCTTTTGTTCAGATGCC 2835

Db 1415 YYYYYYYYYYYY 1427
Qy 1282 CGGGTGCTCCTC 1294

RESULT 8 I66494 7218 bp DNA PAT 23-DEC-1997
LOCUS Sequence 14 from patent US 5670367.
DEFINITION I66494
ACCESSION g2724471
NID
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7218)
AUTHORS Dörner, F., Scheiflinger, F. and Falkner, F. Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES Location/Qualifiers
source
1. 7218
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN

Query Match 1.3%; Score 49; DB 21; Length 7218;
Best Local Similarity 2.0%; Pred. NO. 1.39e-13;
Matches 6; Conservative 165; Mismatches 122; Indels 0; Gaps 0;

Db 1150 YYYYYYYYYYYY 1209
Cp 2080 CCGCGCGGCTGAGTGGACGACGCTGACAGTGCCTTCCCTCGAGGAGACGCTC 2021
Db 1210 YYYYYYYYYYYY 1269
Cp 2020 GGCCCTCTTCTCTCGGACGCTTCTGGCTCCACGACGAGTGCATGTTCAATCGG 1961
Db 1270 YYYYYYYYYYYY 1329
Cp 1960 CCGAGCCGCTGAGTGGGATGAAGCGGAGTCTGACGCTCAGCAGCGGCGGCTGGC 1901
Db 1330 YYYYYYYYYYYY 1389
Cp 1900 TTCCGATGCGCTGACCTGCTTCCGAGCTCCGACGCTCCGACACTCTTCCAGTG 1841
Db 1390 YYYYYYYYYYYY 1442
Cp 1840 CTGCTGATTCATGCTTTCACACTTCTCCACACTCTTCCGGTAGAAA 1788

RESULT 9
LOCUS AC005369 74371 bp DNA PRI 01-AUG-1998
DEFINITION Homo sapiens chromosome 5, BAC clone 119j3 (LBNL H175), complete sequence.
ACCESSION AC005369
NID g3367505
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 74371)
AUTHORS Kimmerly, W., Bondoc, M., Cheng, J., Connolly, K.S., Gunning, K.M.,
Rojeski, H., Subramanian, S. and Martin, C.H.

TITLE Sequencing of human chromosome 5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 74371)
AUTHORS Ricke, D.O.
TITLE Large Scale Sequence Analysis and Annotation with the Sequence
Comparison Analysis (SCAN) System
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 74371)

AUTHORS Kimmerly, W., Bondoc, M., Cheng, J., Connolly, K.S., Gunning, K.M.,
Davis, C.A., Kadner, K., Miguel, T., Pittluck, S., Pollard, M.,
Rojeski, H., Subramanian, S. and Martin, C.H.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
COMMENT Sequence submitted by:
DOE Joint Genome Institute.
FEATURES Location/Qualifiers
source
1. 74371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="5q"
/clone="119j3"
/chromosome="5"
/note="LBNL H175"
893. 1030
/rpt_family="Alu"
repeat_region
2295. 2438
/rpt_family="Alu"
repeat_region
2818. 2859
/note="(GT)21"
/rpt_type=tandem
/rpt_unit=GT
misc_feature
join(3246..3410,3721..3828)
/standard_name="RLF"
/note="65% & 69% protein identity GenPept:U02377"
3431. 3724
/rpt_family="Alu"
repeat_region
3707. 3728
/note="(A)22"
/rpt_type=tandem
/rpt_unit=A
repeat_region
4366. 4661
/rpt_family="Alu"
repeat_region
5327. 5602
/rpt_family="Alu"
repeat_region
6586. 6956
/rpt_family="L1"
repeat_region
6647. 6684
/note="(CA)19"
/rpt_type=tandem
/rpt_unit=CA
repeat_region
7113. 7373
/rpt_family="Alu"
complement(7830..8185)
/standard_name="possible repeat"
8258. 8503
/rpt_family="Alu"
repeat_region
9070. 9387
/rpt_family="Alu"
repeat_region
complement(9740..9845)
/rpt_family="MER42"
repeat_region
complement(10440..11015)
/rpt_family="Alu"
repeat_region
11950. 12250
/rpt_family="Alu"
repeat_region
12057. 12085
/note="(A)29"
/rpt_type=tandem
/rpt_unit=A
repeat_region
12365. 12645
/rpt_family="Alu"
repeat_region
13727. 13750
/note="(AC)12"
/rpt_type=tandem
/rpt_unit=AC
repeat_region
13783. 14024
/rpt_family="L1"
repeat_region
14175. 14470
/rpt_family="Alu"
complement(14905..15259)
repeat_region

QY 131 CTGCCGCTGCCACGCTCGCGCGCTGGCGCCAGCGCTGGCGCTGGTGGCAGCGC 190
Db 143 SRNMGASVGSNTRKTHAKNSADGKVGSKNNGDRNNRYGTGTGKSNVNNCGGGRDVS 202
QY 191 GGGGACCGCGCGCTTCCGCGCGCTGGTGGCCAGTGCCTGTGTGCGCGCTGGGAC 250
Db 203 SYANNKCGSSC 214
QY 251 GCACGCGCGCC 262

RESULT 11
LOCUS 128278 215 bp DNA PAT 30-OCT-1996
DEFINITION Sequence 5 from patent US 5569830.
ACCESSION 128278
NID g1819054
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 215)
AUTHORS Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
TITLE Plant inhibitors of fungal polygalacturonases and their use to control fungal disease
JOURNAL Patent: US 5569830-A 5 29-OCT-1996;
FEATURES Location/Qualifiers
source 1..215
BASE COUNT 15 a 8 c 25 g 26 t 141 others
ORIGIN
Query Match 0.98; Score 35; DB 21; Length 215;
Best Local Similarity 13.28; Pred. No. 8.17e-05;
Matches 26; Conservative 76; Mismatches 95; Indels 0; Gaps 0;

Db 6 MSSSVVSRASCNDRAKKDGTSSHTTDCNRTWCVCTDTTYRVNNDGHNKYSAN 65
Cp 1616 AGTCCCGCAGCTCATCTCCACGTACGCTCCTCGACGAGAGCTTGGCATGCTCCCA 1557
Db 66 YNTGGNNGVGAATHYTHTVNSGADSTVTDYNSAGTSSNGTDCNRSAGDSYSSKT 125
Cp 1556 GGAGATGAACTTCTTGCTTCTGAGGAGCGGCTTGTGTGCTGGAGCCCGCAGA 1497
Db 126 AMFSRNRTKTANNVADSVGSDKNTKTHAKNSADGKVGSKNNGDRNNRYGTGT 185
Cp 1496 GGCTGGGGCAGCAGCGCGCGCAGCGCAGCGCCGACGAGCGGTACACCTGCAGGGC 1437
Db 186 KSNVNNCGGGRDVS 202
Cp 1436 TGCTGTGCTGGCGAGC 1420

RESULT 12
LOCUS AC005369 74371 bp DNA PRI 01-AUG-1998
DEFINITION Homo sapiens chromosome 5, BAC clone 119j3 (LBNL H175), complete sequence.
ACCESSION AC005369
NID g3367505
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 74371)
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE Sequencing of human chromosome 5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 74371)
AUTHORS Ricke,D.O.
TITLE Large Scale Sequence Analysis and Annotation with the Sequence

Comparison Analysis (SCAN) System
Unpublished
3 (bases 1 to 74371)
Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.
Direct Submission
Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, MS 74-157, Berkeley, CA 94720, U.S.A.
Sequence submitted by:
DOE Joint Genome Institute.
FEATURES Location/Qualifiers
source 1..74371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="5q"
/clone="119j3"
/chromosome="5"
/note="LBNL H175"
893..1030
/rpt_family="Alu"
2295..2438
/rpt_family="Alu"
2818..2859
/note="(GT)21"
/rpt_type=tandem
/rpt_unit=GT
join(3246..3410,3721..3828)
/standard_name="RLF"
/note="65% & 69% protein identity GenPept:U22377"
3431..3724
/rpt_family="Alu"
3707..3728
/note="(A)22"
/rpt_type=tandem
/rpt_unit=A
4366..4661
/rpt_family="Alu"
5327..5602
/rpt_family="Alu"
6586..6956
/rpt_family="L1"
6647..6684
/note="(CA)19"
/rpt_type=tandem
/rpt_unit=CA
7113..7373
/rpt_family="Alu"
complement(7830..8185)
/standard_name="possible repeat"
8258..8503
/rpt_family="Alu"
9070..9387
/rpt_family="Alu"
complement(9740..9845)
/rpt_family="MER42"
complement(10440..11015)
/rpt_family="Alu"
11950..12250
/rpt_family="Alu"
12057..12085
/note="(A)29"
/rpt_type=tandem
/rpt_unit=A
12365..12645
/rpt_family="Alu"
13727..13750
/note="(AC)12"
/rpt_type=tandem
/rpt_unit=AC
13783..14024
/rpt_family="L1"

| | | | | | | | | | | | | | | | | | | | |
|-----------------------|-----|---|--|--------|--|--------------------------|--|------|--|-------------|--|--|--|--|--|--|--|--|--|
| | | | | | | LTAQAAVRRRAAREIRIPLYKKG" | | | | | | | | | | | | | |
| BASE COUNT | | 16 a | | 66 c | | 51 g | | 22 t | | 46 others | | | | | | | | | |
| ORIGIN | | | | | | | | | | | | | | | | | | | |
| Query Match | | 0.8%; Score 30; DB 21; Length 201; | | | | | | | | | | | | | | | | | |
| Best Local Similarity | | 36.18; Pred. No. 5.30e-02; | | | | | | | | | | | | | | | | | |
| Matches | | 52; Conservative 35; Mismatches 57; Indels 0; Gaps 0; | | | | | | | | | | | | | | | | | |
| <hr/> | | | | | | | | | | | | | | | | | | | |
| Db | 2 | TGCCCACTCCSGTSTCSCCGAGTCSCCGSCGCSACGCCGCGSCGSCGSCGSC 61 | | | | | | | | | | | | | | | | | |
| QY | 27 | TGCTGGCACGTGGGAAGCCCTGGCCCCGCCACCCC CGCGATGCCGCGCTCCCGCT 86 | | | | | | | | | | | | | | | | | |
| Db | 62 | CSTGSTSTCSCGSCGCTSCTSGAGGGSGGSGCGSGTSGTSGGGGCCTSGCSTST 121 | | | | | | | | | | | | | | | | | |
| QY | 87 | GCCGAGCGTGCGCTCCCTCTGCGACGACTACCGCGAGGTGCTGCCGCTGCCACGT 146 | | | | | | | | | | | | | | | | | |
| Db | 122 | CSGCTSCSCGCTSCACGGCCAGG 145 | | | | | | | | | | | | | | | | | |
| QY | 147 | TCGTGGCGCGCTGGGCCCCAGG 170 | | | | | | | | | | | | | | | | | |
| <hr/> | | | | | | | | | | | | | | | | | | | |
| RESULT 14 | | | | | | | | | | | | | | | | | | | |
| LOCUS | | A10158 | | 201 bp | | DNA | | PAT | | 01-DEC-1993 | | | | | | | | | |
| DEFINITION | | S.griseus gene for preprotease (partial). | | | | | | | | | | | | | | | | | |
| ACCESSION | | A10158 | | | | | | | | | | | | | | | | | |
| NID | | 9490196 | | | | | | | | | | | | | | | | | |
| KEYWORDS | | Streptomyces griseus. | | | | | | | | | | | | | | | | | |
| SOURCE | | Streptomyces griseus | | | | | | | | | | | | | | | | | |
| ORGANISM | | Eubacteria; Firmicutes; Actinomycetes; Streptomycetes; | | | | | | | | | | | | | | | | | |
| REFERENCE | | Streptomycetaceae; Streptomyces. | | | | | | | | | | | | | | | | | |
| AUTHORS | | 1 (bases 1 to 201) | | | | | | | | | | | | | | | | | |
| TITLE | | Garvin,R.T. and James,E. | | | | | | | | | | | | | | | | | |
| JOURNAL | | Production of active proteins containing cystine residues | | | | | | | | | | | | | | | | | |
| PATENT: | | EP 0222279 A 2 20-MAY-1987; | | | | | | | | | | | | | | | | | |
| FEATURES | | Cangene Corporation | | | | | | | | | | | | | | | | | |
| source | | Location/Qualifiers | | | | | | | | | | | | | | | | | |
| CDS | | 1..201 | | | | | | | | | | | | | | | | | |
| | | /organism="Streptomyces griseus" | | | | | | | | | | | | | | | | | |
| | | /db_xref="taxon:1911" | | | | | | | | | | | | | | | | | |
| | | 1..>201 | | | | | | | | | | | | | | | | | |
| | | /codon_start=1 | | | | | | | | | | | | | | | | | |
| | | /transl_table=11 | | | | | | | | | | | | | | | | | |
| | | /product="preprotease" | | | | | | | | | | | | | | | | | |
| | | /db_xref="PID:g490197" | | | | | | | | | | | | | | | | | |
| | | /translation="MPHSVPSPAESPAQPGRPRPVVSRLLGGRAVLGALALSASP | | | | | | | | | | | | | | | | | |
| BASE COUNT | | 13 a | | 66 c | | 54 g | | 22 t | | 46 others | | | | | | | | | |
| ORIGIN | | | | | | | | | | | | | | | | | | | |
| Query Match | | 0.8%; Score 30; DB 21; Length 201; | | | | | | | | | | | | | | | | | |
| Best Local Similarity | | 36.18; Pred. No. 5.30e-02; | | | | | | | | | | | | | | | | | |
| Matches | | 52; Conservative 35; Mismatches 57; Indels 0; Gaps 0; | | | | | | | | | | | | | | | | | |
| <hr/> | | | | | | | | | | | | | | | | | | | |
| Db | 2 | TGCCCACTCCSGTSTCSCCGAGTCSCCGSCGCSACGCCGCGSCGSCGSCGSCGSC 61 | | | | | | | | | | | | | | | | | |
| QY | 27 | TGCTGGCACGTGGGAAGCCCTGGCCCCGCCACCCC CGCGATGCCGCGCTCCCGCT 86 | | | | | | | | | | | | | | | | | |
| Db | 62 | CSTGSTSTCSCGSCGCTSCTSGAGGGSGGSGCGSGTSGTSGGGGCCTSGCSTST 121 | | | | | | | | | | | | | | | | | |
| QY | 87 | GCCGAGCGTGCGCTCCCTCTGCGACGACTACCGCGAGGTGCTGCCGCTGCCACGT 146 | | | | | | | | | | | | | | | | | |
| Db | 122 | CSGCTSCSCGCTSCACGGCCAGG 145 | | | | | | | | | | | | | | | | | |
| QY | 147 | TCGTGGCGCGCTGGGCCCCAGG 170 | | | | | | | | | | | | | | | | | |
| <hr/> | | | | | | | | | | | | | | | | | | | |
| RESULT 15 | | | | | | | | | | | | | | | | | | | |
| LOCUS | | A10162 | | 201 bp | | DNA | | PAT | | 01-DEC-1993 | | | | | | | | | |
| DEFINITION | | Synthetic DNA for preprotease leader & prochymosin. | | | | | | | | | | | | | | | | | |
| ACCESSION | | A10162 | | | | | | | | | | | | | | | | | |
| NID | | 9489115 | | | | | | | | | | | | | | | | | |
| KEYWORDS | | | | | | | | | | | | | | | | | | | |

THIS PAGE BLANK (USPTO)

WILEY

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 01:59:35 1998; MasPar time 10944.33 Seconds
Tabular output not generated. 621.452 Million cell updates/sec

Title: >US-08-951-733-19
Description: (1-3798) from US08951733.seq
Perfect Score: 3798
N.A. Sequence: 1 CCACGGGTCCGGCAGCGCT.....GGAATAGTCATCCCTGAT 3798
Comp: GGTGCGCAGGCCGTCGCGA.....CCTATCAGGTAGGCGACTA

Scoring table: TABLE default
Gap open 30; Gap extend 1
Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-est55
Database: genbank-est107
5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est20 17:gb_est21
18:gb_est21 19:gb_est23 20:gb_est4 21:gb_est5 22:gb_est6
23:gb_est7 24:gb_est8 25:gb_est9 26:gb_gss1 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 12.047; Variance 2.654; scale 4.539

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|-------------|-------------------------|
| 1 | 385 | 10.1 | 389 | 8 | AA281296 | EST08902.r1 NCI_CGAP_GC |
| 2 | 248 | 6.5 | 409 | 24 | AA311750 | EST182469 Jurkat T-cel |
| 3 | 82 | 2.2 | 315 | 13 | AA748707 | nv02e05.s1 NCI_CGAP_GC |
| 4 | 84 | 2.2 | 340 | 13 | AA811084 | oa85c05.s1 NCI_CGAP_GC |
| 5 | 58 | 1.5 | 252 | 12 | AA754459 | 97SN1787 Rice Immature |
| 6 | 58 | 1.5 | 252 | 12 | AA754459 | 97SN1787 Rice Immature |
| 7 | 52 | 1.4 | 303 | 24 | AA299878 | EST12462 Uterus tumor |
| 8 | 49 | 1.3 | 247 | 12 | AA754458 | 97SN1784 Rice Immature |
| 9 | 47 | 1.2 | 247 | 12 | AA754458 | 97SN1784 Rice Immature |
| 10 | 33 | 0.9 | 375 | 23 | AA200728 | mul3h09.r1 Soares 2NbM |
| 11 | 33 | 0.9 | 2275 | 11 | AF034173 | Homo sapiens ntcon2 co |
| 12 | 31 | 0.8 | 660 | 11 | AF034177 | Homo sapiens ntcon6 co |

| | | | | | | | | |
|---|----|----|-----|------|----|----------|------------------------|----------|
| C | 13 | 25 | 0.7 | 143 | 18 | AI171394 | EST217350 Normalized r | 1.41e-03 |
| | 14 | 25 | 0.7 | 181 | 6 | AA386387 | EST81369 Prostate glan | 1.41e-03 |
| | 15 | 25 | 0.7 | 187 | 13 | AA855630 | vw68h11.r1 Stratagene | 1.41e-03 |
| | 16 | 25 | 0.7 | 213 | 16 | R72797 | Y109c09.r1 Homo sapien | 1.41e-03 |
| C | 17 | 25 | 0.7 | 242 | 15 | AI012107 | EST206558 Normalized r | 1.41e-03 |
| C | 18 | 25 | 0.7 | 278 | 13 | AA882818 | TENS0393 T. cruzi epim | 1.41e-03 |
| | 19 | 25 | 0.7 | 301 | 14 | AA848961 | EST191723 Normalized r | 1.41e-03 |
| C | 20 | 25 | 0.7 | 302 | 15 | AI010844 | EST205295 Normalized r | 1.41e-03 |
| | 21 | 27 | 0.7 | 317 | 27 | AQ008007 | CIT-HSP-228C3.TF CIT- | 1.41e-03 |
| | 22 | 25 | 0.7 | 318 | 14 | AA799774 | EST189271 Normalized r | 1.41e-03 |
| | 23 | 25 | 0.7 | 318 | 16 | H50134 | vo27a07.r1 Homo sapien | 1.41e-03 |
| | 24 | 25 | 0.7 | 343 | 22 | AA098755 | T3985 MVAT4 bloodstrea | 1.19e-04 |
| | 25 | 25 | 0.7 | 359 | 24 | AA303595 | EST16305 Aorta endothe | 1.41e-03 |
| | 26 | 25 | 0.7 | 382 | 5 | R54656 | Y74404.r1 Homo sapien | 1.41e-03 |
| | 27 | 26 | 0.7 | 396 | 19 | T26788 | T529 Trypanosoma bruce | 1.19e-04 |
| | 28 | 26 | 0.7 | 400 | 16 | H4707 | Y24c08.r1 Homo sapien | 1.19e-04 |
| C | 29 | 25 | 0.7 | 420 | 17 | AI071447 | UI-R-C1-ku-c-05-0-UI.s | 1.41e-03 |
| C | 30 | 25 | 0.7 | 432 | 6 | AA043978 | 2458f01.s1 Soares preg | 1.41e-03 |
| C | 31 | 25 | 0.7 | 438 | 14 | AA924206 | UI-R-Al-dx-a-11-0-UI.s | 1.41e-03 |
| C | 32 | 25 | 0.7 | 478 | 14 | AA925372 | UI-R-Al-ee-d-07-0-UI.s | 1.41e-03 |
| C | 33 | 25 | 0.7 | 487 | 16 | R86860 | Ym86a03.r1 Homo sapien | 1.41e-03 |
| C | 34 | 25 | 0.7 | 504 | 15 | AI010120 | EST204571 Normalized r | 1.41e-03 |
| C | 35 | 25 | 0.7 | 511 | 17 | AI072026 | UI-R-C2-nd-e-12-0-UI.s | 1.41e-03 |
| C | 36 | 25 | 0.7 | 511 | 15 | AA997735 | UI-R-C0-hn-b-08-0-UI.s | 1.41e-03 |
| C | 37 | 25 | 0.7 | 540 | 14 | AA901048 | UI-R-Al-dv-e-09-0-UI.s | 1.41e-03 |
| | 38 | 27 | 0.7 | 560 | 12 | AA622227 | ac79e03.s1 Stratagene | 1.41e-06 |
| | 39 | 25 | 0.7 | 590 | 10 | AA622227 | nc43c05.s1 NCI_CGAP.Pr | 1.41e-03 |
| C | 40 | 27 | 0.7 | 660 | 11 | AF034177 | Homo sapiens ntcon6 co | 9.41e-06 |
| C | 41 | 25 | 0.7 | 677 | 18 | AI105428 | EST214717 Normalized r | 1.41e-03 |
| C | 42 | 25 | 0.7 | 705 | 14 | AA941159 | L25143.5prime LD dros | 1.41e-03 |
| | 43 | 28 | 0.7 | 1287 | 12 | AF038250 | Homo sapiens clone ntc | 6.90e-07 |
| C | 44 | 28 | 0.7 | 1287 | 12 | AF038250 | Homo sapiens clone ntc | 6.90e-07 |
| C | 45 | 26 | 0.7 | 2275 | 11 | AF034173 | Homo sapiens ntcon2 co | 1.19e-04 |

ALIGNMENTS

RESULT 1
LOCUS AA281296 389 bp mRNA EST 14-AUG-1997
DEFINITION nt08902.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:712562 5',
mRNA sequence.
ACCESSION AA281296
NID 91924194
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 389)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TUMOR National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 2187 Std Error: 0.00
Seq primer: -28ml3 rev2 Et from Amersham
High quality sequence stop: 385.
Location/Qualifiers
1. 389
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer

FEATURES

source

[5'-TGTTACCAATCTCAAGTGGAGCGCGGCTCAATTTTCTTTTCTTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo." /db_xref="taxon:9606" /clone="IMAGE:11562" /clone_lib="NCI_CGAP_GCB1" /tissue_type="germinal center B cell" /lab_host="DH10B" <1. ->389 87 a 102 c 123 g 77 t

BASE COUNT
ORIGIN

Query Match 10.1%; Score 385; DB 8; Length 389;
Best Local Similarity 99.5%; Pred. No. 0.00e+00;
Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 GCCAAGTTCCTGCTGCTGATGAGTGTGACGTCGTGAGCTGCTCAGTCTTTCTTT 60
QY 1691 GCCAAGTTCCTGCTGCTGATGAGTGTGACGTCGTGAGCTGCTCAGTCTTTCTTT 1750

Db 61 TATGTCAGGAGACACCTTTCAAGACAGGCTCTTTTCTACCGAAGAGTCTCTGG 120
QY 1751 TATGTCAGGAGACACCTTTCAAGACAGGCTCTTTTCTACCGAAGAGTCTCTGG 1810

Db 121 AGCAAGTTGCAAGCATTTGGAATCAGACAGACCTTGAAGAGGTTGCAGCTCGGGAGCTG 180
QY 1811 AGCAAGTTGCAAGCATTTGGAATCAGACAGACCTTGAAGAGGTTGCAGCTCGGGAGCTG 1870

Db 181 TCGGAAGCAGAGGTCAGGAGCATCGGGAAGCCAGCCGCCCTGCTGACGTCCAGACTC 240
QY 1871 TCGGAAGCAGAGGTCAGGAGCATCGGGAAGCCAGCCGCCCTGCTGACGTCCAGACTC 1930

Db 241 CGCTTCATCCCAAGCTGACGGGTCGGCCGATTTGTAATGATGATAGTCTGCTGGGA 300
QY 1931 CGCTTCATCCCAAGCTGACGGGTCGGCCGATTTGTAATGATGATAGTCTGCTGGGA 1990

Db 301 GCCAAGCTTCCGAGAAAGAGGCGGAGCGCTCTCACTCGAGGTTGAAGGCACTG 360
QY 1991 GCCAAGCTTCCGAGAAAGAGGCGGAGCGCTCTCACTCGAGGTTGAAGGCACTG 2050

Db 361 TTCACGGTGTCAATACGAGGGGCGCG 389
QY 2051 TTCACGGTGTCAATACGAGGGGCGCG 2079

RESULT 2 A311750 409 bp mRNA EST 19-APR-1997
LOCUS EST182469 Jurkat T-cells VI Homo sapiens cDNA 5' end, mRNA
DEFINITION sequence.
ACCESSION A311750
NID g1964077
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae;
Homo.
REFERENCE 1 (bases 1 to 409)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.B., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,

He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl.), 3-174 (1995)
96026280

Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES
source
1..409
/organism="Homo sapiens"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
/db_xref="ATCC (inhost):158964"
/db_xref="taxon:9606"
/clone_lib="Jurkat T-cells VI"
/cell_type="T-lymphocyte"
BASE COUNT 65 a 120 c 133 g 86 t 5 others
ORIGIN
Query Match 6.5%; Score 248; DB 24; Length 409;
Best Local Similarity 99.2%; Pred. No. 0.00e+00;
Matches 248; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 GTTTGGTGGATGATTCTTGTGTGACACCTCACCTCACCCACCGGAAACCTTCCTCA 60
QY 2661 GTTTGGTGGATGATTCTTGTGTGACACCTCACCTCACCCACCGGAAACCTTCCTCA 2720

Db 61 GGACCTTGGTCCGAGGTGTCCTGAGTATGGTGGTGAACCTTCGGAAGACAGTGG 120
QY 2721 GGACCTTGGTCCGAGGTGTCCTGAGTATGGTGGTGAACCTTCGGAAGACAGTGG 2780

Db 121 TGAACCTTCCCTGTAGAGACGAGCCCTGGTGCACGGCTTTTTCAGATCCCGGCC 180
QY 2781 TGAACCTTCCCTGTAGAGACGAGCCCTGGTGCACGGCTTTTTCAGATCCCGGCC 2840

Db 181 ACGGCTATTCCTGCTGGCGCTGCTGCTGGATACCCGACCTCGAGGTGCAGAGCG 240
QY 2841 ACGGCTATTCCTGCTGGCGCTGCTGCTGGATACCCGACCTCGAGGTGCAGAGCG 2900

Db 241 ACTACTCCAG 250
QY 2901 ACTACTCCAG 2910

RESULT 3 AA748707 315 bp mRNA EST 18-FEB-1998
LOCUS ny02e05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1270592,
DEFINITION mRNA sequence.
ACCESSION AA748707
NID g2788665
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 315)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index
Unpublished (1997)

JOURNAL
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1226 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 281.

FEATURES
source
1. .315
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer
[5'-GCTTACCAATCTGAAGTGGAGCGCGCCCTCATTTT-TTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

BASE COUNT 68 a 95 c 87 g 65 t
ORIGIN
Query Match 2.2%; Score 82; DB 13; Length 315;
Best Local Similarity 98.8%; Pred. No. 1.90e-88;
Matches 83; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 231 ATCTGGGATGGACTATTCCTATGTGGGAGTGAAGCGCGGCTCCTGGTGAGGAAAGC 290
Cp 3798 ATCAGGGATGGACTATTCCTATGTGGGAGTGAAGCGCGGCTCCTGGTGAGGAAAGC 3739
Db 291 TGGCCCTGGGTGGAGCGCGGC 314
Cp 3738 TGGCCCTGGGTGGAGCGCGGC 3715

RESULT 4
LOCUS AA811084 340 bp mRNA EST 19-FEB-1998
DEFINITION ca85c05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1319048, mRNA sequence.
ACCESSION AA811084
NID g2880695
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 340)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 2249 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 331.

FEATURES
source
1. .340
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer
[5'-GCTTACCAATCTGAAGTGGAGCGCGCCCTCATTTT-TTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

BASE COUNT 73 a 86 c 100 g 81 t
ORIGIN

Query Match 2.2%; Score 84; DB 13; Length 340;
Best Local Similarity 98.8%; Pred. No. 7.88e-92;
Matches 85; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 240 ATCTGGGATGGACTATTCCTATGTGGGAGTGAAGCGCGGCTCCTGGTGAGGAAAGC 299
Cp 3798 ATCAGGGATGGACTATTCCTATGTGGGAGTGAAGCGCGGCTCCTGGTGAGGAAAGC 3739
Db 300 TGGCCCTGGGTGGAGCGCGGC 325
Cp 3738 TGGCCCTGGGTGGAGCGCGGC 3713

RESULT 5
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII CDNA Library Oryza sativa cDNA clone 97SN1787, mRNA sequence.
ACCESSION AA754459
NID g2801165
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
REFERENCE 1 (bases 1 to 252)
AUTHORS Nahn, B.H., Kim, J.K., Cheong, J.J., Kim, S.I., Hahn, T.R., Moon, E.P., Kim, W.T., Kim, W.Y., Yang, M.S., Park, R.D., Sohn, U.I., Kang, K.Y., Lee, M.C. and Eun, M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT Contact: Eun M.Y.

Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.astil.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bnhnm@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
Location/Qualifiers

FEATURES

source

1. .252
/organism="Oryza sativa"
/cultivar="Milyang23"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"

BASE COUNT 5 a 21 c 12 g 35 t 179 others

ORIGIN

Query Match 1.5%; Score 58; DB 12; Length 252;
Best Local Similarity 10.0%; Pred. No. 3.51e-49;
Matches 25; Conservative 120; Mismatches 105; Indels 0; Gaps 0;

Db 1 HWDCTMTVRCGCCBAMMTTSYBCHGNBVVWVASHGNVMSVHNCBTRGTHCDCKKNVW 60
QY 3032 TTGAGGTGAACAGCTCCACAGCGGTGGACCAACATCTACAGATCCTCTGTCGAG 3091
Db 61 STMTGTVMNBVSGDWHYWBVNTKVDVGNHTRCSRWRBYTRMAHYHDYTNCBYNNND 120
QY 3092 CGGTACAGGTTTCACGATGTCGTCAGCTCCCATTTTCATCAGCAAGTTTGAAGAAC 3151
Db 121 YMHMBBYBTGCTMTCWCBYHNTKCTASGWHSTNYDVKSTNTWGTBVSYDYSMH 180
QY 3152 CCCACATTTTCTCGCGGTCTATCTGCACAGCGGCTCCCTGCTACTCCTCGA 3211
Db 181 GYWCBBYKHYTKYSTTRATSYTCVRKYCVMMTKKVVKKYHVVBGCHTDSKCKTMW 240
QY 3212 GCCAAGACGAGGAGTCTGCTGGGCGCAAGGCGCGCCCTCTGCCCTCCGAG 3271
Db 241 MTNRHVMFTST 250
QY 3272 GCGGTGCAGT 3281

RESULT 6 AA754459 252 bp mRNA EST 20-JAN-1998
LOCUS 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
DEFINITION cDNA clone 97SN1787, mRNA sequence.
ACCESSION AA754459
NID 92801165
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.

REFERENCE 1 (bases 1 to 252)
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.Y., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.

TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT Contact: Eun M.Y.
Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea

Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.astil.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bnhnm@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
Location/Qualifiers

FEATURES

source

1. .252
/organism="Oryza sativa"
/cultivar="Milyang23"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"

BASE COUNT 5 a 21 c 12 g 35 t 179 others

ORIGIN

Query Match 1.5%; Score 58; DB 12; Length 252;
Best Local Similarity 12.8%; Pred. No. 3.51e-49;
Matches 23; Conservative 88; Mismatches 69; Indels 0; Gaps 0;
Db 24 BCHGNBVVWVASHGNVMSVHNCBTRGTHCDCKKNVMSVHNCBTRGTHCDCKKNVW 83
Cp 894 CAGGTGACACACACACAGACAGGTCACCTCGCTCCACGCGTCTGCTCGCGGTCGCC 835
Db 84 NTKVDVGNHTRCSRWRBYTRMAHYHDYTNCBYNNNDYHMHBBYBTTGCTCTMTCWB 143
Cp 834 CAGGACCCCTGCCACAGCGGTCGCTCGCTCAGGCGCAGCGCACGCTCGGCTC 775
Db 144 HYNTKCTASGWHSTNYDVKSTNTWGTBVSYDKSMHYWCBBYKHYTKYSTTRATSY 203
Cp 774 TTGGCAACGCGAGACTTCGCTGGCAGTCCCGCCGCTCTCTCGCACCCGCGGCTGCG 715

RESULT 7 AA299878 303 bp mRNA EST 18-APR-1997
LOCUS EST12462 Uterus tumor I Homo sapiens cDNA 5' end, mRNA sequence.
DEFINITION AA299878
ACCESSION 91952209
NID
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.

REFERENCE

AUTHORS

1 (bases 1 to 303)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitchugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palanco,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.

TITLE

JOURNAL

MEDLINE

COMMENT

96026280

[illegible]

| | | | | |
|----|-----|---|---|----|
| Qy | 183 | TC | AGTCGGGGGACCTGGCGGGCTTTCGGCGCTGTGTGGGCCCCAGTGCCTGGTGTGGCTGC | 24 |
| Db | 88 | TNTGGTGKGTTVNVHSGNNRCSNVVYVMBYAYCDYBHYBDRANHVDDTRCTNDRGYC | 147 | |
| Qy | 243 | CCTGGGACGACGGCGCGCCCCCGCGCGCCCTCCTTCGCCCGAGTGCTCTGCCTGAAGG | 302 | |
| Db | 148 | NYTASDNGTSAIRVTVGYDKTDSDCGGGCKWRKVYTGSSBYBRGCVNVVWRTTSMWTDKST | 207 | |
| Qy | 303 | AGCTGTGTGGCCGAGTCTCGACGAGGCTGTGCGAGCGCGCGCGGAGAACGCTGCTGGCCT | 362 | |

363 TCGGCTTCGGCTGCTGGACGGGCCCG 390

| | | | | | |
|------------|--|--------|------|-----|-------------|
| LOCUS | AA734356 | 247 bp | mRNA | ESI | 20-JAN-1998 |
| DEFINITION | Oryza sativa | | | | |
| | Rice Immature Seed Lambda ZAPII cDNA Library | | | | |
| | cdna clone 97SN1784, mRNA sequence. | | | | |

| SOURCE | ORGANISM |
|-----------|---|
| ESI. | <i>Oryza sativa</i> |
| REIWKURUS | Eukaryotae; Viridiplantae; Charophyta/Embryophyta group; |
| | Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida |
| | Poales; Poaceae; <i>Oryza</i> . |

AUTHORS Nahm, B. H., Kim, J. K., Cheong, J. J., Kim, S. I., Hahn, T. R., Moon, E. P., Kim, W. T., Kim, W. Y., Yang, M. S., Park, R. D., Sohn, U. I., Kang, K. Y., Lee, N. C. and Eun, M. Y.

TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed

JOURNAL Unpublished (1998)

Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307

FEATURES

Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji University, Yongin, Korea. 449-728 bhnahm@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
Location/Qualifiers

```

/clone="97SN1784"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
BASE COUNT      7 a   16 c   21 g   34 t   169 others
ORIGIN
Query Match      1.3%   Score 47;   DB 12;   Length 247;
Best local similarity 4.88;   Pred NO 2 25e-32;

```

| | | | | |
|-----------------------|--------|---------------------|--------|-------------|
| Query Match | 1.28; | Score 47; | DB 12; | Length 247; |
| Best Local Similarity | 14.88; | Pred. No. 2.25e-32; | | |

[illegible]

| LOCUS | 187 bp | EST | 06-MAR-1998 |
|------------|------------|-----------------------|-----------------------------|
| DEFINITION | Stratagene | mouse heart (#937316) | Mus musculus cDNA clone |
| | v60811.r1 | TR:Q99960 | Q99960 PLAKOPHILIN 2A.[1] ; |
| | 1260165.5 | | mRNA |

sequence-
ACCESSION AA355630
NID 92943168

EST. house mouse.
SOURCE Mus musculus
ORGANISM

REFERENCE
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 187)
Marta M Hillier I Allen M Bowles M Dietrich N Dubucq F

AUTHORS
Marras, M.S., Hillier, L.L., Allread, M., Bowles, M., Dietrich, N., Dubouque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schelleng, B., K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Marras, M.S.

TITLE
The WASRU-HHMI Mouse EST Project
Unpublished (1996)

COMMENT: Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel.: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information
Fax: 314 286 1810

MG1:66217
 Inverse Consortium (info@genome.mg1.hawaii.gov) for further information.
 Seq primer: -28ml3 rev1 ET from Amersham
 High quality sequence stop: 178.

| FEATURES | source |
|---------------------|--------------------------|
| Location/Qualifiers | 1..187 |
| | /organism="Mus musculus" |

```

/strain="NIH/Swiss"
/note="Organ: heart; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
old_36_2"

```

Oligo dt. 93 pooled NIH/swiss 13 day embryo hearts.
Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5,
adaptor sequence: 5', GAATTCGGCAGCAG 3', -3', adaptor
sequence: 5', CTCGACATTTTTTTTTTTTTTTT 3',

```
sequence: 5 CACGACIIIIIIIIIIIIIIIIII 3'
/db_xref="taxon:10090"
/clone="1260185"
/clone_lib="Stratagene mouse heart (#937316)"
```

```

/stratigene model= heart (W37316)
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"

```

```

BASE COUNT      29 a      65 c      70 g      23 t
ORIGIN
/lab_host="SOLR (kanamycin resistant)"

```

Query Match 0.7% Score 25; DB 13; Length 187;
Best Local Similarity 85.7% Pred. No. 1,416-03;
Matches 30; Conservative 0; Microtubules 5; Reddy 0;

| Matches | 30; | Conservative | 0; | Mismatches | 5; | Indels | 0; | Gaps | 0; |
|---------|-----|--------------------------------|-----|------------|----|--------|----|------|----|
| Db | 105 | GGTGGCGGCAGCAGCGCGCGGCACCGCGGC | 139 | | | | | | |

Qy 170 GGCIGCGCGCTGGTGACGCGGGGACCCGGCGGC 204

Search completed: Fri Dec 25 05:48:10 1998
Job time : 13715 secs.

that the expressed peptide contains 2 or 4 cysteine residues positioned in, or flanking, the unpredicted or variant residues. These residues confer some degree of conformational rigidity to the peptides. The TSARS or compans. comprising a TSAR binding domain can be used *in vivo* to deliver a chemically or biologically active moiety, eg. metal ion, radioisotope, peptide, toxin or enzyme, to the specific target or on the cell. They can also replace the function of macromolecules, eg. monoclonal or polyclonal antibodies and therefore circumvent the need for complex methods of hybridoma formation or *in vivo* antibody production. The TSARS are easily characterised and have designed activity allowing direct and rapid detection in a screening process.

| | | | | |
|-----------------------|-------|---------------------|--------|----------------|
| Query Match | 1.0%; | Score 38; | DB 12; | Length 114; |
| Best Local Similarity | 4.5%; | Pred. NO. 1.25e-06; | | |
| Matches | 5; | Conservative | 34; | Mismatches 73; |
| | | | Indels | 0; |
| | | | Gaps | 0; |

QY 212 GCGCTGGTGGCCAGTGCCTGGTGTGGTGTGGACGACGGCGCCC 263

```

KW erector coman; concatenated heterofunctional protein; linker;
KW direct; concatenated heterofunctional protein; linker;
KW rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key
FT Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
FT FT

```

PA (UINC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
DR P-PSDB; R65150 and R65151.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure: Page 35: 255pp. English.

RB5151-54. TSARs are concatenated heterofunctional proteins or peptides, comprising at least two functional regions - a binding domain with affinity for a ligand and a second effector peptide portion that is chemically or biologically active. They may further comprise a linker peptide between the 2 domains. The oligonucleotides are also designed so that the expressed peptide contains 2 or 4 cysteine residues positioned in, or flanking, the unpredicted or variant residues. These residues confer some degree of conformational rigidity to the peptides. The TSARs or comps. comprising a TSAR binding domain can be used *in vivo* to deliver a chemically or biologically active moiety, eg. metal ion, radioisotope, peptide, toxin or enzyme, to the specific target or on the cell. They can also replace the function of macromolecules, eg.

16. **Comments:** _____

Db 63 bnbnnbnnbnnbnnbgtcnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
 Cp 1483 CAGCCGGCAGGAGCCGCGACGAGCGCTACACTGCAGGGGTGTG 1432

RESULT 12

ID Q70467 standard; DNA; 114 BP.
 AC Q70467;
 DT 05-APR-1995 (first entry)
 DE Generic DNA sequence to generate a random TSAR peptide library.
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; generic; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 55..60
 FT /*tag= a
 FT /note= "this sequence represents 'Z'; Z can be a
 FT sequence of 6, 9 or 12 nucleotides (see
 FT comments)"
 PN W09418318-A.
 PD 18-AUG-1994. U00977.
 PF 01-FEB-1994; US-013416.
 PR 01-FEB-1993; US-176500.
 PR 30-DEC-1993; US-176500.
 PR 31-JAN-1994; US-189331.
 PA (UYNC-) UNIV NORTH CAROLINA.
 PI Fowlkes DM, Kay BK;
 DR WPI: 94-279739/34.
 DR P-PSDB: R65153.
 DT Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PT comprising a binding domain and an effector domain
 PS Disclosure; Page 35; 255pp; English.
 CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be
 CC represented as follows: X(NNB)16(TGC)(NNB)12(TGC)(NNB)11(Y). X
 CC and Y are flanking restriction sites (X is not the same as Y) that are
 CC not specified further. Other generic sequences are shown in Q70466-68.
 CC Other specific peptides generated by these generic sequences are shown in
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
 CC comprising at least two functional regions - a binding domain with
 CC affinity for a ligand and a second effector peptide portion that is
 CC chemically or biologically active. They may further comprise a linker
 CC peptide between the 2 domains. The oligonucleotides are also designed so
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned
 CC in, or flanking, the unpredicted or variant residues. These residues
 CC confer some degree of conformational rigidity to the peptides. The TSARs
 CC or compsns. comprising a TSAR binding domain can be used in vivo to
 CC deliver a chemically or biologically active moiety, eg. metal ion,
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
 CC cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need for
 CC complex methods of hybridoma formation or in vivo antibody production.
 CC The TSARs are easily characterised and have designed activity allowing
 CC direct and rapid detection in a screening process.
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.0%; Score 38; DB 12; Length 114;
 Best Local Similarity 4.5%; Pred. No. 1.25e-06;
 Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;

Db 3 bnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 62
 Cp 1558 CAGGAGATGAACCTCTGTGTCCTCTGAGGAGCGCGTTCGTGTGCGTGGAGCCCA 1499
 Db 63 bnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
 Cp 1498 GAGCGCTGGGGCCACGAGCGCGCGAGCGAGCGCGCCGACGACGCGTACAC 1447

RESULT 13

ID Q70469 standard; DNA; 114 BP.

AC Q70469;
 DT 07-APR-1995 (first entry)
 DE Generic DNA sequence to generate a random TSAR peptide library.
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; generic; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 55..60
 FT /*tag= a
 FT /note= "this sequence represents 'Z'; Z can be a
 FT sequence of 6, 9 or 12 nucleotides (see
 FT comments)"
 PN W09418318-A.
 PD 18-AUG-1994. U00977.
 PF 01-FEB-1994; US-013416.
 PR 01-FEB-1993; US-176500.
 PR 30-DEC-1993; US-176500.
 PR 31-JAN-1994; US-189331.
 PA (UYNC-) UNIV NORTH CAROLINA.
 PI Fowlkes DM, Kay BK;
 DR WPI: 94-279739/34.
 DT Identifying proteins or peptide(s) which bind a ligand - by
 PT screening a recombinant vector library expressing fusion proteins
 PT comprising a binding domain and an effector domain
 PS Disclosure; Page 35; 255pp; English.
 CC Q70469 is a generic DNA sequence used to generate random TSAR peptide
 CC This generic formula can be represented as follows: X(TGC)(NNB)10-
 CC (TGC)(NNB)62(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
 CC sites (X is not the same as Y) that are not specified further. This
 CC sequence generates peptides that are cloverleaf in structure. Other
 CC generic sequences are shown in Q70465-68. Other specific peptides
 CC generated by these generic sequences are shown in R65150-54. TSARs are
 CC concatenated heterofunctional proteins or peptides, comprising at least
 CC two functional regions - a binding domain with affinity for a ligand and
 CC a second effector peptide portion that is chemically or biologically
 CC active. They may further comprise a linker peptide between the 2 domains.
 CC The oligonucleotides are also designed so that the expressed peptide
 CC contains 2 or 4 cysteine residues positioned in, or flanking, the
 CC unpredicted or variant residues. These residues confer some degree of
 CC conformational rigidity to the peptides. The TSARs or compsns. comprising
 CC a TSAR binding domain can be used in vivo to deliver a chemically or
 CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
 CC or enzyme, to the specific target or on the cell. They can also replace
 CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
 CC and therefore circumvent the need for complex methods of hybridoma
 CC formation or in vivo antibody production. The TSARs are easily
 CC characterised and have designed activity allowing direct and rapid
 CC detection in a screening process.
 SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;
 Query Match 1.0%; Score 37; DB 12; Length 114;
 Best Local Similarity 6.3%; Pred. No. 4.11e-06;
 Matches 7; Conservative 32; Mismatches 72; Indels 0; Gaps 0;
 Db 2 gcnbn 61
 Cp 3572 GCGGTGCGGGCTGGGTGGGGCGCGCCCTCCTCCTCGGAGCGCGGTGAC 3513
 Db 62 nbnnbtgcnbt 112
 Cp 3512 AGGCTGCTGTCTGCTCTCGGCTGTGCGCGGTGCGGCGCATCAGT 3462
 RESULT 14
 ID T76405 standard; DNA; 178 BP.
 AC T76405;
 DT 15-SEP-1997 (first entry)
 DE Human endothelin-1 antisenese oligonucleotide.
 KW Asthma; airway epithelium; adenosine free; cystic fibrosis;
 KW chronic obstructive pulmonary disease; bronchitis; ss.
 OS Synthetic.
 PN W09640162-A1.

| | |
|----|---|
| PD | 19-DEC-1996. |
| PF | 06-JUN-1996; UO9306. |
| PR | 07-JUN-1995; US-474497. |
| PA | (UYEC-) UNIV EAST CAROLINA. |
| PI | Metzger WJ, Nyce JW; |
| DR | WPL: 97-051871/05. |
| PT | Treatment of airway diseases such as asthma - by topically applying |
| PT | adenosine-free antisense oligo:nucleotide to airway epithelium of |
| PS | subject |
| PS | Claim 38; Page 38; 71pp; English. |
| CC | A method for treating airway disease in a subject has been produced, |
| CC | which involves the topical administration of an essentially adenosine |
| CC | free antisense oligonucleotide (ON) to the airway epithelium of the |
| CC | subject. The present sequence is an antisense oligonucleotide specific |
| CC | for the human endothelin-1, targeted at the initiation codon. The |
| CC | method can be used to treat airway diseases such as cystic fibrosis, |
| CC | asthma, chronic obstructive pulmonary disease, bronchitis and other |
| CC | airway diseases characterised by an inflammatory response. By |
| CC | eliminating adenosine from the antisense ON, its liberation upon |
| CC | antisense degradation is prevented, thereby preventing adenosine- |
| CC | induced bronchoconstriction in patients with hyper-reactive airways. |
| SQ | Sequence 178 BP; 0 A; 52 C; 46 G; 32 T; |
| | Query Match 1.0%; Score 39; DB 32; Length 178; |
| | Best Local Similarity 30.1%; Pred. No. 3.78e-07; |
| | Matches 22; Conservative 34; Mismatches 17; Indels 0; Gaps 0 |
| Dd | 105 cbgbcbbcbcbgbgbbgtcgtbgcbbbttcccttgcggbbbbbbgggbt 164 : : : : : : : : : : : : : : : |
| Qy | 197 CCGCGCGCTTCGCGCGCTGGTGCCACAGTGCCTGTGTGCTGCCCTGGGACGCACGG 256 : : : : : : : : : : : : : : : |
| Dd | 165 cbbbbbectcccq 177 : : : : : : : : : : : : : : : |
| Qy | 257 CCGCCCCCGCGC 269 : : : : : : : : : : : : : : : |

RESULT 15
ID T76405 standard; DNA; 178 BP.
AC T76405;
DT 15-SEP-1997 (first entry)
DE Human endothelin-1 antisense oligonucleotide.
KW Asthma; airway epithelium; adenosis free; cystic fibrosis;
KW chronic obstructive pulmonary disease; bronchitis; ss.
OS Synthetic.
PN WO9640162-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; UO9306.
PR 07-JUN-1995; US-474497.
PR (OYEC-) UNIV EAST CAROLINA.
PI Metzger WJ, Nyce JW;
WPI: 97-051871/05.
DR Treatment of airway diseases such as asthma - by topically applying
PT adenosine-free antisense oligo:nucleotide to airway epithelium of
PT subject
PS Claim 5; Page 38; 71pp: English.
CC A method for treating airway disease in a subject has been produced,
CC which involves the topical administration of an essentially adenosine
CC free antisense oligonucleotide (ON) to the airway epithelium of the
CC subject. The present sequence is an antisense oligonucleotide specific
CC for the human endothelin-1, targeted at the initiation codon. The
CC method can be used to treat airway diseases such as cystic fibrosis,
CC asthma, chronic obstructive pulmonary disease, bronchitis and other
CC airway diseases characterised by an inflammatory response. By
CC eliminating adenosis from the antisense ON, its liberation upon
CC antisense degradation is prevented, thereby preventing adenosis-
CC induced bronchoconstriction in patients with hyper-reactive airways.
SQ Sequence 178 BP; 0 A; 52 C; 46 G; 32 T;

Query Match 1.4%; Score 38; DB 32; Length 178;
Best Local Similarity 31.4%; Pred. No. 1.25e-06;
Matches 37; Conservative 41; Mismatches 40; Indels 0; Gaps 0

Db 52 cbcccgcgcgctgbgcctcgccttctgctgtcttctggbgcctctggcggcc 111

```
Cp 3585 CCTCAGACTCCACGGCTCGGSGCTGTGGCGCCGCCCTCCCTCTGGACGTAG 3526  
      |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  
Db 112 cbbcbcgbggbgbbbcbtcbtgcbttcgtbbbbbgbgcbtcbbbb 169  
      :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Cp 3525 AGCCCGCGTGACAGGCGTCTGCCTCTCGGCTGGCTGGCGGGGTGCC 3468
```

Search completed: Fri Dec 25 06:17:23 1998
Job time : 1733 secs.

Search completed: Fri Dec 25 06:17:23 1998
Job time : 1733 secs.

THIS PAGE BLANK (USPTO)

MPsrch_nna n.a. - n.a. database search, using Smith-Waterman algorithm
with gap extension penalties as well as gap opening penalties

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Run on: Fri Dec 25 06:26:25 1998; MasPar time 6934.12 Seconds
Tabular output not generated. 586.919 Million cell updates/sec

Title: >US-08-951-733-19
Description: (1-3798) from US08951733.seq
Perfect Score: 3798
N.A. Sequence: 1 CCACGCGTCCGGCGACGCT.....GGAATAGTCCATCCCTGAT 3798
Comp: GGTGCGCAGGCCGTCGCGA.....CCTATCAGGTAGGGGACAT

Scoring table: TABLE default
Gap open 30; Gap extend 1
Nmatch STD : Dbase 0; Query 0
Searched: 1665728 seqs, 535777161 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-pending
1:P9 2:U6000 3:U6001 4:U6002 5:U6003 6:U6004 7:U6005
8:U6006 9:U6007 10:U6008 11:U7 12:U80 13:U81 14:U82
15:U83 16:U84 17:U84B 18:U85 19:U86 20:U87 21:U88
22:U89 23:U90A 24:U90B 25:U91 26:NEWP 27:NEWU6 28:NEWU8
29:NEWU9

Statistics: Mean 11.275; Variance 4.495; scale 2.508
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| SUMMARIES | | | | | |
|------------|-------|-------------|--------|-------------------------------------|-----------|
| Result No. | Score | Query Match | Length | ID | Pred. No. |
| 1 | 3798 | 100.0 | 3798 | 22 US-08-951-Sequence 19, Applicati | 0.00e+00 |
| 2 | 3787 | 99.7 | 7029 | 22 US-08-911-Sequence 1, Applicatio | 0.00e+00 |
| 3 | 3784 | 99.6 | 4015 | 24 US-08-052-Sequence 1, Applicatio | 0.00e+00 |
| 4 | 3784 | 99.6 | 4015 | 22 US-08-912-Sequence 1, Applicatio | 0.00e+00 |
| 5 | 3784 | 99.6 | 4015 | 21 US-08-854-Sequence 224, Applicat | 0.00e+00 |
| 6 | 3781 | 99.6 | 4023 | 23 US-08-026-Sequence 35, Applicati | 0.00e+00 |
| 7 | 3468 | 91.3 | 4029 | 21 US-08-854-Sequence 173, Applicat | 0.00e+00 |
| 8 | 3468 | 91.3 | 4029 | 21 US-08-851-Sequence 173, Applicat | 0.00e+00 |
| 9 | 3391 | 89.3 | 3855 | 22 US-08-912-Sequence 4, Applicatio | 0.00e+00 |
| 10 | 3391 | 89.3 | 3855 | 22 US-08-911-Sequence 18, Applicati | 0.00e+00 |
| 11 | 2848 | 75.0 | 2848 | 22 US-08-951-Sequence 13, Applicati | 0.00e+00 |
| 12 | 1711 | 45.1 | 2176 | 22 US-08-912-Sequence 3, Applicatio | 0.00e+00 |
| 13 | 1575 | 41.5 | 3346 | 23 US-08-026-Sequence 36, Applicati | 0.00e+00 |
| 14 | 1520 | 40.0 | 2171 | 21 US-08-851-Sequence 100, Applicat | 0.00e+00 |
| 15 | 1520 | 40.0 | 2171 | 21 US-08-854-Sequence 100, Applicat | 0.00e+00 |

| | | | | | | |
|----|------|------|------|----|----------------------------------|----------|
| 16 | 1520 | 40.0 | 2171 | 21 | US-08-846-Sequence 100, Applicat | 0.00e+00 |
| 17 | 1497 | 39.4 | 4200 | 22 | US-08-912-Sequence 6, Applicatio | 0.00e+00 |
| 18 | 1079 | 28.4 | 3496 | 23 | US-09-042-Sequence 1, Applicatio | 0.00e+00 |
| 19 | 949 | 25.0 | 949 | 22 | US-08-951-Sequence 18, Applicati | 0.00e+00 |
| 20 | 535 | 14.1 | 535 | 22 | US-08-911-Sequence 52, Applicati | 0.00e+00 |
| 21 | 460 | 12.1 | 550 | 22 | US-08-911-Sequence 54, Applicati | 0.00e+00 |
| 22 | 385 | 10.1 | 389 | 21 | US-08-846-Sequence 62, Applicati | 0.00e+00 |
| 23 | 385 | 10.1 | 389 | 21 | US-08-844-Sequence 62, Applicati | 0.00e+00 |
| 24 | 385 | 10.1 | 389 | 21 | US-08-854-Sequence 8, Applicatio | 0.00e+00 |
| 25 | 385 | 10.1 | 389 | 22 | US-08-912-Sequence 8, Applicatio | 0.00e+00 |
| 26 | 385 | 10.1 | 389 | 22 | US-08-911-Sequence 17, Applicati | 0.00e+00 |
| 27 | 385 | 10.1 | 389 | 21 | US-08-851-Sequence 62, Applicati | 0.00e+00 |
| 28 | 182 | 4.8 | 182 | 22 | US-08-912-Sequence 9, Applicatio | 2.49e-18 |
| 29 | 120 | 3.2 | 2651 | 23 | US-09-042-Sequence 5, Applicatio | 1.04e-85 |
| 30 | 97 | 2.6 | 240 | 22 | US-08-912-Sequence 7, Applicatio | 3.64e-63 |
| 31 | 83 | 2.2 | 249 | 5 | US-60-035-Sequence 1054, Applic | 9.44e-50 |
| 32 | 78 | 2.1 | 578 | 18 | US-08-569-Sequence 14, Applicati | 4.90e-45 |
| 33 | 78 | 2.1 | 578 | 19 | US-08-624-Sequence 14, Applicati | 4.90e-45 |
| 34 | 69 | 1.8 | 477 | 11 | US-07-904-Sequence 10, Applicati | 1.12e-36 |
| 35 | 69 | 1.8 | 477 | 11 | US-07-904-Sequence 10, Applicati | 1.12e-36 |
| 36 | 68 | 1.8 | 578 | 19 | US-08-624-Sequence 14, Applicati | 9.30e-36 |
| 37 | 68 | 1.8 | 578 | 18 | US-08-569-Sequence 14, Applicati | 9.30e-36 |
| 38 | 69 | 1.8 | 1808 | 23 | US-09-042-Sequence 4, Applicatio | 1.12e-36 |
| 39 | 67 | 1.8 | 7218 | 17 | US-08-466-Sequence 14, Applicati | 7.65e-35 |
| 40 | 66 | 1.7 | 383 | 16 | US-08-446-Sequence 8, Applicatio | 6.25e-34 |
| 41 | 66 | 1.7 | 383 | 15 | US-08-311-Sequence 8, Applicatio | 6.25e-34 |
| 42 | 66 | 1.7 | 383 | 16 | US-08-446-Sequence 8, Applicatio | 6.25e-34 |
| 43 | 66 | 1.7 | 383 | 17 | US-08-451-Sequence 8, Applicatio | 6.25e-34 |
| 44 | 66 | 1.7 | 383 | 16 | US-08-446-Sequence 8, Applicatio | 6.25e-34 |
| 45 | 66 | 1.7 | 383 | 16 | US-08-446-Sequence 8, Applicatio | 6.25e-34 |

ALIGNMENTS

RESULT 1
ID US-08-951-733-19 STANDARD; DNA; UNC; 3798 BP.
AC xxxxxx
DE Sequence 19, Application US/08951733
CC Sequence 19, Application US/08951733
CC GENERAL INFORMATION:
CC APPLICANT: Harrington, Lea A.
CC APPLICANT: Robinson, Murray O.
CC TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS
CC NUMBER OF SEQUENCES: 44
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Amgen Inc.
CC STREET: One Amgen Center Drive
CC CITY: Thousand Oaks
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 91320-1789
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/951,733
CC FILING DATE: 16-OCT-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/873,039
CC FILING DATE: 11-JUN-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/751,189
CC FILING DATE: 15-NOV-1996
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Oleski, Nancy A.
CC REGISTRATION NUMBER: 34,688
CC REFERENCE/DOCKET NUMBER: A-433B
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (805) 447-6504

Db 1981 CGTCGTGGGAGCCAGACGTTCCGACAGAGAAAGAGGCGCGGCGTCTCACCTCGAGGCT 2040
|||||
QY 1981 CGTCGTGGGAGCCAGAACGTTCCGACAGAGAAAGAGGCGCGGCGTCTCACCTCGAGGCT 2040
|||||
Db 2041 GAAGGCACTGTTTCAGCGGTGCTCAACTACAGAGCGGCGCGCGCGGCGTCTCTGGGCGC 2100
|||||
QY 2041 GAAGGCACTGTTTCAGCGGTGCTCAACTACAGAGCGGCGCGCGGCGTCTCTGGGCGC 2100
|||||
Db 2101 CTCTGCTGGGCGCTGGACGATATCCACAGGGCTTGGCGACCTTCGTGCTGGGTGTGG 2160
|||||
QY 2101 CTCTGCTGGGCGCTGGACGATATCCACAGGGCTTGGCGACCTTCGTGCTGGGTGTGG 2160
|||||
Db 2161 GGCCCAAGGACCCCGCTGAGCTGTACTTTGTCAGAGTGGATGTACGGGCGCGTACGA 2220
|||||
QY 2161 GGCCCAAGGACCCCGCTGAGCTGTACTTTGTCAGAGTGGATGTACGGGCGCGTACGA 2220
|||||
Db 2221 CACCATCCCCAGGACAGGCTCACGGAGTCTACGGCAGCATCATCAAAACCCAGAAAC 2280
|||||
QY 2221 CACCATCCCCAGGACAGGCTCACGGAGTCTACGGCAGCATCATCAAAACCCAGAAAC 2280
|||||
Db 2281 GTACTGCGTGGGTGGTATGCCGTGTGTCAGAGGCGGCGCCATGGGACGTCGCGCAAGGC 2340
|||||
QY 2281 GTACTGCGTGGGTGGTATGCCGTGTGTCAGAGGCGGCGCCATGGGACGTCGCGCAAGGC 2340
|||||
Db 2341 CTTTCAGAGCCAGCTCTCTACCTTGACAGACCTCCAGCGGTATCGGACAGTTCGTGGC 2400
|||||
QY 2341 CTTTCAGAGCCAGCTCTCTACCTTGACAGACCTCCAGCGGTATCGGACAGTTCGTGGC 2400
|||||
Db 2401 TCACCTGTCAGGAGACACGCGCTGAGGGATGCCGTGTCATCGAGCAGAGCTCCCTCCCT 2460
|||||
QY 2401 TCACCTGTCAGGAGACACGCGCTGAGGGATGCCGTGTCATCGAGCAGAGCTCCCTCCCT 2460
|||||
Db 2461 GAATGAGGCCAGCAGTGGCTCTTCGACGCTCTTCCTACGCTTCATGTGCCACACGCGCT 2520
|||||
QY 2461 GAATGAGGCCAGCAGTGGCTCTTCGACGCTCTTCCTACGCTTCATGTGCCACACGCGCT 2520
|||||
Db 2521 GCGCATCAGGGGCAAGTCTCTACGTCAGTCGAGGGATCCCGAGGCTCCATCTCTC 2580
|||||
QY 2521 GCGCATCAGGGGCAAGTCTCTACGTCAGTCGAGGGATCCCGAGGCTCCATCTCTC 2580
|||||
Db 2581 CAGCTGCTCTGAGCGCTGTGCTACGCGACATGGAGAAACAGCTGTTTCGGGGATTCG 2640
|||||
QY 2581 CAGCTGCTCTGAGCGCTGTGCTACGCGACATGGAGAAACAGCTGTTTCGGGGATTCG 2640
|||||
Db 2641 GCGGAGCGGCTCTCTCGGTTTGGTGGATGATTTCTTGGTGACACCTCACCTCAC 2700
|||||
QY 2641 GCGGAGCGGCTCTCTCGGTTTGGTGGATGATTTCTTGGTGACACCTCACCTCAC 2700
|||||
Db 2701 CCAGCGGAAAACCTTCCTCAGGACCTGCTCGAGGTGCTCCCTGAGTGGCTGGT 2760
|||||
QY 2701 CCAGCGGAAAACCTTCCTCAGGACCTGCTCGAGGTGCTCCCTGAGTGGCTGGT 2760
|||||
Db 2761 GAACCTGCGGAAGACAGTGTGAACCTTCCTGTAGAGAGAGGCGCTGGGTGGCAGGC 2820
|||||
QY 2761 GAACCTGCGGAAGACAGTGTGAACCTTCCTGTAGAGAGAGGCGCTGGGTGGCAGGC 2820
|||||
Db 2821 TTTTGTTCAGATCCGCGCCACAGCGCTATTCCCTGTGGTGGGCTGTGCTGGATACCCG 2880
|||||
QY 2821 TTTTGTTCAGATCCGCGCCACAGCGCTATTCCCTGTGGTGGGCTGTGCTGGATACCCG 2880
|||||
Db 2881 GACCTGGAGGTGAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCT 2940
|||||
QY 2881 GACCTGGAGGTGAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCT 2940
|||||
Db 2941 CACCTCAACCGGGCTTCAAGGCTGGGAGAACATGCTGCGCAAACTCTTTGGGCTCT 3000
|||||
QY 2941 CACCTCAACCGGGCTTCAAGGCTGGGAGAACATGCTGCGCAAACTCTTTGGGCTCT 3000
|||||
Db 3001 GCGGCTGAAGTGTACAGCGCTGTTTCTGGATTTCAGGTGAACAGCTCCAGAGCGGTG 3060
|||||
QY 3001 GCGGCTGAAGTGTACAGCGCTGTTTCTGGATTTCAGGTGAACAGCTCCAGAGCGGTG 3060
|||||
Db 3061 CACCAACATCTACAAGATCCTCTGCTGACAGCGCTACAGGTTTCACGCAATGTGTGTCGA 3120

QY 3061 CACCAACATCTACAAGATCCTCTGCTGAGGCGTACAGGTTTCACGCAATGTGTGTCGA...3120
|||||
Db 3121 GCTCCCATTTTCATCAGCAAGTTTGAAGAACCCACATTTTCTCTGCGGTCTCATCTCTGA 3180
|||||
QY 3121 GCTCCCATTTTCATCAGCAAGTTTGAAGAACCCACATTTTCTCTGCGGTCTCATCTCTGA 3180
|||||
Db 3181 CAGGGCTCCCTCTGCTACTCTCATCTCTGAAAGCAGAGGATGTCGTGGGGGC 3240
|||||
QY 3181 CAGGGCTCCCTCTGCTACTCTCATCTCTGAAAGCAGAGGATGTCGTGGGGGC 3240
|||||
Db 3241 CAAGGGCGCGCGGCGCTCTGCCCTCCAGGCGCGTGCAGTGGCTGTGCCAACAAGCAT 3300
|||||
QY 3241 CAAGGGCGCGCGGCGCTCTGCCCTCCAGGCGCGTGCAGTGGCTGTGCCAACAAGCAT 3300
|||||
Db 3301 CTTGCTCAAGCTGACTTCGACACCGTGTACCTACGTGCCACCTCTGGGTCTACTCAGGAC 3360
|||||
QY 3301 CTTGCTCAAGCTGACTTCGACACCGTGTACCTACGTGCCACCTCTGGGTCTACTCAGGAC 3360
|||||
Db 3361 AGCCAGAGCGAGTGTGAGTCGGAAGCTCCGGGAGCAGCTGACTGTCCTGGAGGCGC 3420
|||||
QY 3361 AGCCAGAGCGAGTGTGAGTCGGAAGCTCCGGGAGCAGCTGACTGTCCTGGAGGCGC 3420
|||||
Db 3421 AGCCAGCGGCGCTGCTGCTGAGCTTCAAGACCATCTGGACTGATGGCCACCCGCCCA 3480
|||||
QY 3421 AGCCAGCGGCGCTGCTGCTGAGCTTCAAGACCATCTGGACTGATGGCCACCCGCCCA 3480
|||||
Db 3481 CAGCCAGGCGGAGAGCAGACACAGCAGCCCTGTTCACGCGGGCTCTAGCTCCAGGGAG 3540
|||||
QY 3481 CAGCCAGGCGGAGAGCAGACACAGCAGCCCTGTTCACGCGGGCTCTAGCTCCAGGGAG 3540
|||||
Db 3541 GGAGGGCGGCGCCACACCCAGGCGCGCAGCTGGGAGTCTGAGGCTGAGTGTGTTT 3600
|||||
QY 3541 GGAGGGCGGCGCCACACCCAGGCGCGCAGCTGGGAGTCTGAGGCTGAGTGTGTTT 3600
|||||
Db 3601 GCGGAGGCGCTGCTGCTGAGGCTGAGTGTCCGGCTGAGGCTTGGGGAGTGTCTC 3660
|||||
QY 3601 GCGGAGGCGCTGCTGCTGAGGCTGAGTGTCCGGCTGAGGCTTGGGGAGTGTCTC 3660
|||||
Db 3661 CAGCCAAAGGCTGAGTGTCCAGCACACCTGCTCTACTTCCCCACAGGCTGGGGCTC 3720
|||||
QY 3661 CAGCCAAAGGCTGAGTGTCCAGCACACCTGCTCTACTTCCCCACAGGCTGGGGCTC 3720
|||||
Db 3721 GGTCCACCCCGAGGCGAGCTTTTCTCACCAGGAGCGCGCTTCCACTCCCCACATAG 3780
|||||
QY 3721 GGTCCACCCCGAGGCGAGCTTTTCTCACCAGGAGCGCGCTTCCACTCCCCACATAG 3780
|||||
Db 3781 AATAGTCCATCCCTGAT 3798
|||||
QY 3781 AATAGTCCATCCCTGAT 3798
|||||

RESULT 2

ID US-08-911-312-1 STANDARD; DNA; UNC; 7029 BP.

AC xxxxxx

DT

Sequence 1, Application US/08911312

Sequence 1, Application US/08911312

GENERAL INFORMATION:

CC APPLICANT: Cech, Thomas R.

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin B.

CC APPLICANT: Andrews, William

CC TITLE OF INVENTION: Telomerase Reverse Transcriptase

CC NUMBER OF SEQUENCES: 170

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Townsend and Crew LLP

CC STREET: Two Embarcadero Center, Eighth Floor

CC CITY: San Francisco

CC STATE: California

CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/911.312
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Einhorn, Gregory P.
CC REGISTRATION NUMBER: 38,440
CC REFERENCE/DOCKET NUMBER: 015389-002500US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7029 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 782..4177
CC OTHER INFORMATION: /product= "human telomerase reverse
CC OTHER INFORMATION: transcriptase (htrrt)"
CC OTHER INFORMATION: /note= "cDNA contained in plasmid
CC OTHER INFORMATION: pGRN121"
CC SEQUENCE 7029 BP; 1416 A; 2122 C; 2051 G; 1440 T; 0 OTHER.

Query Match 99.7%; Score 3787; DB 22; Length 7029;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 3791; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 718 CGCGTCGACGCGCGCTCCCGCTCCGAGCGGTGCGCTCCCTGCTGCGAGCCACTACCG 837
|||||
Qy 4 CGCGTCGCGGCGCGCTCCCGCTCCGAGCGGTGCGCTCCCTGCTGCGAGCCACTACCG 63
|||||
Db 778 CGCGATCGCGCGCGCTCCCGCTCCGAGCGGTGCGCTCCCTGCTGCGAGCCACTACCG 837
|||||
Qy 64 CGCGATCGCGCGCGCTCCCGCTCCGAGCGGTGCGCTCCCTGCTGCGAGCCACTACCG 123
|||||
Db 838 CGAGGTGCTGCGGTGCGCGACGTTCTGTCGCGCGCTCGGCGCCCGAGGCGTGGCGGTGGT 897
|||||
Qy 124 CGAGGTGCTGCGGTGCGCGACGTTCTGTCGCGCGCTCGGCGCCCGAGGCGTGGCGGTGGT 183
|||||
Db 898 GCAGCGGGGACCGCGCGGTTCGCGCGCGCTGGTGGCGCGCTGGTGGTGGTGGTGGC 957
|||||

Qy 184 GCAGCGCGGGGACCGCGCGCTTTCGCGCGGTGCTGGTGGCCAGTGGTGGTGGTGGC 243
Db 958 CTGGGACGACG 1017
|||||
Qy 244 CTGGGACGACG 303
|||||
Db 1018 GCTGGTGGCGCGAGTGTCTGAGAGGCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCG 1077
|||||
Qy 304 GCTGGTGGCGCGAGTGTCTGAGAGGCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCG 363
|||||
Db 1078 CGGCTTTCGCGCTGCTGAGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137
|||||
Qy 364 CGGCTTTCGCGCTGCTGAGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 423
|||||
Db 1138 GCGCAGCTACTCTGCCCAACACGCTGACCGACGCTGCGCGGGGAGCGGGCGGTGGGGGCT 1197
|||||
Qy 424 GCGCAGCTACTCTGCCCAACACGCTGACCGACGCTGCGGGGAGCGGGCGGTGGGGGCT 483
|||||
Db 1198 GCTGCTGGCGCGGTGGGGAGACGCTGTGTGTACCTGTGTGGTGTGTGTGTGTGTGTGT 1257
|||||
Qy 484 GCTGCTGGCGCGGTGGGGAGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 543
|||||
Db 1258 TGTGCTGGTGGTCCCGAGCTGCGCGCTTACCAGGTGTGCGGCGCGCGCGCTGTACCAGCTCGG 1317
|||||
Qy 544 TGTGCTGGTGGTCCCGAGCTGCGCGCTTACCAGGTGTGCGGCGCGCGCGCTGTACCAGCTCGG 603
|||||
Db 1318 CGCTGCACTACTAGGCG 1377
|||||
Qy 604 CGCTGCACTACTAGGCG 663
|||||
Db 1378 CGAAGCGCGCTTGAACCATAGCTCAGGAGCGCGGGGTGCGCGCGCGCGCGCGCGCGCG 1437
|||||
Qy 664 CGAAGCGCGCTTGAACCATAGCTCAGGAGCGCGGGGTGCGCGCGCGCGCGCGCGCGCG 723
|||||
Db 1438 GGTGCGAGGAGCGCGGGGAGTGCACGAGTGCAGCGAGTGTGCGGTGCGCGCGCGCGCG 1497
|||||
Qy 724 GGTGCGAGGAGCGCGGGGAGTGCACGAGTGCAGCGAGTGTGCGGTGCGCGCGCGCGCG 783
|||||
Db 1498 GCGTGGCGTGGCG 1557
|||||
Qy 784 GCGTGGCGTGGCG 843
|||||
Db 1558 GGCAGAGCGCGTGGAGCGAGTGCACGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1617
|||||
Qy 844 GGCAGAGCGCGTGGAGCGAGTGCACGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 903
|||||
Db 1618 CGAAGAGCGACCTCTTGGAGGTTGGCTCTCTGCGAGCGCGCGCGCGCGCGCGCGCG 1677
|||||
Qy 904 CGAAGAGCGACCTCTTGGAGGTTGGCTCTCTGCGAGCGCGCGCGCGCGCGCGCGCG 963
|||||
Db 1678 GGCAGCGCGAGCG 1737
|||||
Qy 964 GGCAGCGCGAGCG 1023
|||||
Db 1738 GCGTGTGCGCGCGGTGTACCGCGAGACCAAGCACTTCTCTACTCTCTAGGCGCGCAAGGA 1797
|||||
Qy 1024 GCGTGTGCGCGCGGTGTACCGCGAGACCAAGCACTTCTCTACTCTCTAGGCGCGCAAGGA 1083
|||||
Db 1798 GCAAGTGGCGCGCGCTTCTCTACTCTAGTCTCTGAGCGCGCGCGCGCGCGCGCGCG 1857
|||||
Qy 1084 GCAAGTGGCGCGCGCTTCTCTACTCTAGTCTCTGAGCGCGCGCGCGCGCGCGCGCG 1143
|||||
Db 1858 GCTGCTGGAGACCATCTTCTGCGGTTCAGGCGCGCTGGATGCCAGGACTCCCCCGCAGGTT 1917
|||||
Qy 1144 GCTGCTGGAGACCATCTTCTGCGGTTCAGGCGCGCTGGATGCCAGGACTCCCCCGCAGGTT 1203
|||||
Db 1918 GCGCGCGCTGCGCGCGCGTGTGCGCAATGCGCGCGCGCTGTTTCTGTGAGTGTGTGGGAA 1977
|||||
Qy 1204 GCGCGCGCTGCGCGCGCGTGTGCGCAATGCGCGCGCGCTGTTTCTGTGAGTGTGTGGGAA 1263
|||||
Db 1978 CCAGCGCGAGTGGCGCGGTGTCTCTCAAGAGCGACTGCGCGCGCGCTGCGAGCTGGGT 2037
|||||
Qy 1264 CCAGCGCGAGTGGCGCGGTGTCTCTCAAGAGCGACTGCGCGCGCGCTGCGAGCTGGGT 1323
|||||

[illegible]

QY 3484 CCAGGCGGAGGAGACACACAGCAGCCCTGTACCGCGGGCTTACCTCCAGGGAGGA 3543
Db 4258 GGGCGGCGCCACACCCAGGCGCCGACCGCTGGAGTCTGAGGCTGAGTGGTGGC 4317
QY 3544 GGGCGGCGCCACACCCAGGCGCCGACCGCTGGAGTCTGAGGCTGAGTGGTGGC 3603
Db 4318 CGAGGCTGATGTCGGCTGAAGCTGAGTGTCCGCTGAGGCTGAGGCTGAGTGGCAG 4377
QY 3604 CGAGGCTGATGTCGGCTGAAGCTGAGTGTCCGCTGAGGCTGAGGCTGAGTGGCAG 3663
Db 4378 CCAAGGCTGAGTGTCCAGCACACCTCGCTTTCACCTCCACAGGCTGGGCTGGC 4437
QY 3664 CCAAGGCTGAGTGTCCAGCACACCTCGCTTTCACCTCCACAGGCTGGGCTGGC 3723
Db 4438 TCCACCCAGGCGGAGCTTTTCTCCACAGGAGCGCGCTTCCACTCCACATAGGAAT 4497
QY 3724 TCCACCCAGGCGGAGCTTTTCTCCACAGGAGCGCGCTTCCACTCCACATAGGAAT 3783
Db 4498 AGTCCATCCCGAT 4512
QY 3784 AGTCCATCCCGAT 3798

RESULT 3
ID US-09-052-919-1 STANDARD; DNA; UNC; 4015 BP.
AC xxxxxx
DE
DT
Sequence 1, Application US/09052919
CC Sequence 1, Application US/09052919
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Antisense Compositions for Detecting and
CC TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase
CC NUMBER OF SEQUENCES: 72
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/052,919
CC FILING DATE: 31-MAR-1998
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/974,549
CC FILING DATE: 19-NOV-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/974,584
CC FILING DATE: 19-NOV-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Parent, Annette S.
CC REGISTRATION NUMBER: 42,058
CC REFERENCE/DOCKET NUMBER: 015389-003600US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4015 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 56..3454
CC OTHER INFORMATION: /product= "human telomerase reverse
CC OTHER INFORMATION: transcriptase (hTERT)"
CC SEQUENCE 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 OTHER.
SQ

Query Match 99.6%; Score 3784; DB 24; Length 4015;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 3785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 1 GCAGCGCTGCTGCTGCTGCGACGCTGGGAAGCCCTGGCGCCGACCCCGCGATGCC 60
QY 13 GCAGCGCTGCTGCTGCTGCGACGCTGGGAAGCCCTGGCGCCGACCCCGCGATGCC 72
Db 61 GCGCGCTCCCGCTGCGGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGGAGGTGCT 120
QY 73 GCGCGCTCCCGCTGCGGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGGAGGTGCT 132
Db 121 GCGCGTGGCAGGTTGCTGCGGCGCTGGGGCCCGCAGGCGCTGGCGCTGGTGCAGCGCG 180
QY 133 GCGCGTGGCAGGTTGCTGCGGCGCTGGGGCCCGCAGGCGCTGGCGCTGGTGCAGCGCG 192
Db 181 GGACCGCGCGGCTTTCCGCGCGCTGGTGCCCGAGTCCCTGCTGCGCTCCCTGGGACGC 240
QY 193 GGACCGCGCGGCTTTCCGCGCGCTGGTGCCCGAGTCCCTGCTGCTGCTGGGACGC 252
Db 241 ACGCGCGCGCGCGCGCGCGCTCCCTTCCTCCCGCAGGTGCTTCCCTGAAGGAGTGGTGGC 300
QY 253 ACGCGCGCGCGCGCGCGCGCTCCCTTCCTCCCGCAGGTGCTTCCCTGAAGGAGTGGTGGC 312
Db 301 CCGAGTGTGCGAGAGGCTGTGCGAGCGCGCGCGGCGGAGAACGTCGTGGCTTCGCTTCGC 360
QY 313 CCGAGTGTGCGAGAGGCTGTGCGAGCGCGCGCGGCGGAGAACGTCGTGGCTTCGCTTCGC 372
Db 361 GCTGCTGGAGCGGCG 420
QY 373 GCTGCTGGAGCGGCG 432

QY 2593 CAGCGTGTCTACGGCGACATGGAGAAACAGCTGTTTCGGGGGATTCCGGCGGACGGGCT 2652
Db 2641 GCTCTCGGTTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCACGGGAAAC 2700
QY 2653 GCTCTCGGTTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCACGGGAAAC 2712
Db 2701 CTTCTCAGGACCTGGTCCGAGGTGTCCTGAGTATGGCTGCGTGGTGAACCTGCGGAA 2760
QY 2713 CTTCTCAGGACCTGGTCCGAGGTGTCCTGAGTATGGCTGCGTGGTGAACCTGCGGAA 2772
Db 2761 GACAGTGTGAACCTTCCCTGTAGAGAGAGGCGCTTGGTGGACAGGCTTGTGTCAGAT 2820
QY 2773 GACAGTGTGAACCTTCCCTGTAGAGAGAGGCGCTTGGTGGACAGGCTTGTGTCAGAT 2832
Db 2821 GCGCGCCACAGGCTTATCCCTGTGTGGGCGCTGCTGTGATACCCGAGACCTTGGAGGT 2880
QY 2833 GCGCGCCACAGGCTTATCCCTGTGTGGGCGCTGCTGTGATACCCGAGACCTTGGAGGT 2892
Db 2881 GCAGAGCCACTTCCAGCTATGCCGAGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940
QY 2893 GCAGAGCCACTTCCAGCTATGCCGAGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2952
Db 2941 CGGCTTCAAGGCTGGAGGAAACATCGCTGCAAACTCTTTGGGCTTGGCGCTGAAGTG 3000
QY 2953 CGGCTTCAAGGCTGGAGGAAACATCGCTGCAAACTCTTTGGGCTTGGCGCTGAAGTG 3012
Db 3001 TCACAGCCTGTTTTCGAGTGTGAGGTAACAGCCTCCAGAGCTGTGCACCAACATCTA 3060
QY 3013 TCACAGCCTGTTTTCGAGTGTGAGGTAACAGCCTCCAGAGCTGTGCACCAACATCTA 3072
Db 3061 CAAGATCTCTGCTGCAGGCGTACAGTTCACGCATGTGTGAGCTCCCATTTCA 3120
QY 3073 CAAGATCTCTGCTGCAGGCGTACAGTTCACGCATGTGTGAGCTCCCATTTCA 3132
Db 3121 TCAGCAAGCTTGGAGAAACCCACATTTTCTGCGCTGCTCTGACAGCGCTCCCT 3180
QY 3133 TCAGCAAGCTTGGAGAAACCCACATTTTCTGCGCTGCTCTGACAGCGCTCCCT 3192
Db 3181 CTGCTACTCCATCTCTGAAAGCAAGACGAGGATGTCTGAGGGGCAAGGGCGCGC 3240
QY 3193 CTGCTACTCCATCTCTGAAAGCAAGACGAGGATGTCTGAGGGGCAAGGGCGCGC 3252
Db 3241 CGGCGCTCTGCGCTCCGAGGCGGTGAGTGTGCTGCTGACCAAGCATTTCTGCTCAAGCT 3300
QY 3253 CGGCGCTCTGCGCTCCGAGGCGGTGAGTGTGCTGCTGACCAAGCATTTCTGCTCAAGCT 3312
Db 3301 GACTCGACACCGTGTACCTAGTGGCTACTCTGGGCTACTCAGGACAGCCAGAGCA 3360
QY 3313 GACTCGACACCGTGTACCTAGTGGCTACTCTGGGCTACTCAGGACAGCCAGAGCA 3372
Db 3361 GCTGAGTCGGAAGCTTCCCGGGGACAGCTGACTGCCCTGGAGGCGCGAGCAACCGCGC 3420
QY 3373 GCTGAGTCGGAAGCTTCCCGGGGACAGCTGACTGCCCTGGAGGCGCGAGCAACCGCGC 3432
Db 3421 ACTGCCCTCAGACTTCAGACATCTCGACTGATGGCCACCGCCACAGCCAGGCGGA 3480
QY 3433 ACTGCCCTCAGACTTCAGACATCTCGACTGATGGCCACCGCCACAGCCAGGCGGA 3492
Db 3481 GAGCAGACACAGCAGCCTGTGACGCGGCTCTACCTCCAGGAGGAGGGCGGCGC 3540
QY 3493 GAGCAGACACAGCAGCCTGTGACGCGGCTCTACCTCCAGGAGGAGGGCGGCGC 3552
Db 3541 CACACCCAGGCGCCAGCGCTGGAGTCTGAGGCTGAGTGTGTCGCGGAGGCGCTG 3600
QY 3553 CACACCCAGGCGCCAGCGCTGGAGTCTGAGGCTGAGTGTGTCGCGGAGGCGCTG 3612
Db 3601 CATGTCGCGCTGAGGCTGAGTGTCCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3660
QY 3613 CATGTCGCGCTGAGGCTGAGTGTCCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3672
Db 3661 GAGTGTCCAGCACACTCTCGCTTCTTACTTCCACAGGCTGGCGCTGGCTCCACCCCA 3720
QY 3673 GAGTGTCCAGCACACTCTCGCTTCTTACTTCCACAGGCTGGCGCTGGCTCCACCCCA 3732

Db 3721 GGGCCAGCTTTTCTCTACCAGGAGCGCGGCTTCCACTCCCCACATAGATAGTCCATCC 3780
QY 3733 GGGCCAGCTTTTCTCTACCAGGAGCGCGGCTTCCACTCCCCACATAGATAGTCCATCC 3792
Db 3781 CCAGAT 3786
QY 3793 CCTGAT 3798
RESULT 4
ID US-08-912-951-1 STANDARD; DNA; UNC; 4015 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/08912951
CC Sequence 1, Application US/08912951
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC TITLE OF INVENTION: THERAPEUTIC METHODS
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002600US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4015 base pairs

1933 CTTATCCCAAGCTGACGGGCTCGGCCGATGTGAAATGACTACGTCTGTGGAGC 1992
1981 CAGAACCTTCGCGAGAGAAAAGAGGCCGAGCGTCTACCTCGAGGGTGAAGGCACCTGTT 2040
1993 CAGAACCTTCGCGAGAGAAAAGAGGCCGAGCGTCTACCTCGAGGGTGAAGGCACCTGTT 2052
2041 CAGCGTGTCTAACTACGAGGCGGCGGCGGCCCGCCGCTCTCTGGCGGCTCTGTGCTGGG 2100
2053 CAGCGTGTCTAACTACGAGGCGGCGGCGGCCCGCCGCTCTCTGGCGGCTCTGTGCTGGG 2112
2101 CCGTACCATATCCACAGGGCTGCGCACCTCTGCTGTGCTGTGCGGGGCCAGGACCC 2160
2113 CCGTACCATATCCACAGGGCTGCGCACCTCTGCTGTGCTGTGCGGGGCCAGGACCC 2172
2161 GCGCGCTGAGGTGTACTTTGTCAAGTGGATGTGACGGGCGCGTACGACACCATCCCCCA 2220
2173 GCGCGCTGAGGTGTACTTTGTCAAGTGGATGTGACGGGCGCGTACGACACCATCCCCCA 2232
2221 GGACAGGCTCAGGAGGTCTACGCGAGCATCATCAAAACCCAGAACAGTACTGGGTGG 2280
2233 GGACAGGCTCAGGAGGTCTACGCGAGCATCATCAAAACCCAGAACAGTACTGGGTGG 2292
2281 TCGGTATCGCGTGTCCAGAGGCCCGCCATGGCGCACCTCCGCAAGGCTTCAAGAGCCA 2340
2293 TCGGTATCGCGTGTCCAGAGGCCCGCCATGGCGCACCTCCGCAAGGCTTCAAGAGCCA 2352
2341 CGTCTCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTACCTGCGGA 2400
2353 CGTCTCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTACCTGCGGA 2412
2401 GACACGCGGCTGAGGATGCGTGTCTATCGAGCAGAGCTCTCCCTGATGAGGCCAG 2460
2413 GACACGCGGCTGAGGATGCGTGTCTATCGAGCAGAGCTCTCCCTGATGAGGCCAG 2472
2461 CAGTGGCTCTTCGACGCTTCTTACGCTTCATGTGCGCACCGCCGCGCATCAGGG 2520
2473 CAGTGGCTCTTCGACGCTTCTTACGCTTCATGTGCGCACCGCCGCGCATCAGGG 2532
2521 CAACTCTACGCTTCCAGTCCAGGGGATCCCGAGGGCTCCATCTCTCCAGGCTGCTGTG 2580
2533 CAACTCTACGCTTCCAGTCCAGGGGATCCCGAGGGCTCCATCTCTCCAGGCTGCTGTG 2592
2581 CAGCGTGTCTACGCGGACATGGAGAACAGCTGTTTGGGGGATTCGGGGGAGCGGCT 2640
2593 CAGCGTGTCTACGCGGACATGGAGAACAGCTGTTTGGGGGATTCGGGGGAGCGGCT 2652
2641 GCTCTCGCTTGTGGTGGATTTCTTCTGTGTGACACCTCACCTCACCCAGCGAANAAC 2700
2653 GCTCTCGCTTGTGGTGGATTTCTTGTGTGACACCTCACCTCACCCAGCGAANAAC 2712
2701 CTTCTCTCAGGACCTTGTGCGGAGGTGCTCCCTGAGTATGGCTGCGTGTGAACCTTGGGAA 2760
2713 CTTCTCTCAGGACCTTGTGCGGAGGTGCTCCCTGAGTATGGCTGCGTGTGAACCTTGGGAA 2772
2761 GACAGTGTGAATCTTCCCTGTAGAGACAGGCGCTCGGTCGAGCGGCTTTGTTTCAGAT 2820
2773 GACAGTGTGAATCTTCCCTGTAGAGACAGGCGCTCGGTCGAGCGGCTTTGTTTCAGAT 2832
2821 GCGGCGCCACGCGCTATTCCCTGTGCGGCTGCTGTGATACCCGAGCCCTGGAGGT 2880
2833 GCGGCGCCACGCGCTATTCCCTGTGCGGCTGCTGTGATACCCGAGCCCTGGAGGT 2892
2881 GCAGAGCGACTACTCCAGTATGCGCGGACCTTCCATCAGAGCCAGTCTACCTTCAACCG 2940
2893 GCAGAGCGACTACTCCAGTATGCGCGGACCTTCCATCAGAGCCAGTCTACCTTCAACCG 2952
2941 GCGCTTCAAGGCTGGAGAGACATCGCTCGCAAACTCTTTGGGGTCTTTGGCGCTGAAGTG 3000
2953 GCGCTTCAAGGCTGGAGAGACATCGCTCGCAAACTCTTTGGGGTCTTTGGCGCTGAAGTG 3012
3001 TCACAGCGCTGTTTCTGATTTGACGTTGAACAGCGCTCCAGAGGCTGACCAACATCTA 3060

3013 TCACAGCGCTGTTTCTGATTTGCAGGTGAACAGGCTCCAGACGGTGTGCACCAACATCTA 3072
3061 CAAGATCCTCTGCTGCGAGGCGTACAGGTTTTCACGCAATGTGTGCTGAGCTCCCATTTCA 3120
3073 CAAGATCCTCTGCTGCGAGGCGTACAGGTTTTCACGCAATGTGTGCTGAGCTCCCATTTCA 3132
3121 TCACAGGTTTGGAAAGCCACATTTTCTGCGGGTCTATCTCTACAGCGGCTCCCT 3180
3133 TCACAGGTTTGGAAAGCCACATTTTCTGCGGGTCTATCTCTGACAGCGGCTCCCT 3192
3181 CTGCTACTCTCATCTGAAAGCCAAAGACGAGGATGTGCTGGGGCCCAAGGCGCGCG 3240
3193 CTGCTACTCTCATCTGAAAGCCAAAGACGAGGATGTGCTGGGGCCCAAGGCGCGCG 3252
3241 CGGCGCTCTGCGCTCCGAGGCGCGTGCAGTGGCTGTGCCACCAAGCATTTCTGCTCAAGCT 3300
3253 CGGCGCTCTGCGCTCCGAGGCGCGTGCAGTGGCTGTGCCACCAAGCATTTCTGCTCAAGCT 3312
3301 GACTGACACCGTGTACCTACGTGCCACTCTCTGGGGTCTACTCAGGACAGCCACAGCGCA 3360
3313 GACTGACACCGTGTACCTACGTGCCACTCTCTGGGGTCTACTCAGGACAGCCACAGCGCA 3372
3361 GCTGAGTCGGAAGCTCCCGGGGACGCGTGCAGTCCCTTGAGGCCGCGACCAACCGCGC 3420
3373 GCTGAGTCGGAAGCTCCCGGGGACGCGTGCAGTCCCTTGAGGCCGCGACCAACCGCGC 3432
3421 ACTGCCCTCAGACTTCAAGACCATCTCTGAGTGTGATGCCACCCGCGCCACAGCCAGGCGCA 3480
3433 ACTGCCCTCAGACTTCAAGACCATCTCTGAGTGTGATGCCACCCGCGCCACAGCCAGGCGCA 3492
3481 GAGCAGACACAGGACCGCTGTACGCGGGGCTCTAGTCCAGGGAGGAGGGGCGCGC 3540
3493 GAGCAGACACAGGACCGCTGTACGCGGGGCTCTACGCGGGGCTCTAGTCCAGGGAGGCGCGC 3552
3541 CACACCCAGGCGCGCGCTGGGAGTGTGAGGCTCAGTGTGAGTGTGCGCGAGGCGCTG 3600
3553 CACACCCAGGCGCGCGCTGGGAGTGTGAGGCTCAGTGTGAGTGTGCGCGAGGCGCTG 3612
3601 CATGTCGCGCTGAGGCTGAGTGTGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3660
3613 CATGTCGCGCTGAGGCTGAGTGTGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3672
3661 GAGTGTCCAGCACACCTGCGGCTCTTCACTTCCCGACAGGCTGCGGCTGCGGCTCCACCCCA 3720
3673 GAGTGTCCAGCACACCTGCGGCTCTTCACTTCCCGACAGGCTGCGGCTGCGGCTCCACCCCA 3732
3721 GGGCCAGCTTTTCTCTACCAGGACCGCGCTTCCACTCCCGACATAGTAATAGTCCATCC 3780
3733 GGGCCAGCTTTTCTCTACCAGGACCGCGCTTCCACTCCCGACATAGTAATAGTCCATCC 3792
3781 CCAGAT 3786
3793 CCTGAT 3798

RESULT 5

ID US-08-854-050-224 STANDARD; DNA; UNC; 4015 BP.
AC xxxxxx

Sequence 224, Application US/08854050

Sequence 224, Application US/08854050

GENERAL INFORMATION:

CC APPLICANT: Cech, Thomas R.

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin

CC APPLICANT: Andrews, William H.

CC TITLE OF INVENTION: Novel Telomerase

CC NUMBER OF SEQUENCES: 225

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: DS/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536

CC OTHER INFORMATION: component
SQ SEQUENCE 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 OTHER.

| | | | | |
|-----------------------|----------------|--------------------|----------|-------------|
| Query Match | 99.6% | Score 3784 | DB 21 | Length 4015 |
| Best Local Similarity | 100.0% | Pred. No. 0.00e+00 | | |
| Matches 3785 | Conservative 0 | Mismatches 1 | Indels 0 | Gaps 0 |

| | | | |
|----|-----|--|-----|
| Db | 1 | GCACGCTGGCTCTCTGTCGCGACGTTGGGAAGCCCTGGCCCCGGGCACACCCCGCATGCC | 60 |
| | | | |
| Qy | 13 | GCACGCTGGCTCTCTGTCGCGACGTTGGGAAGCCCTGGCCCCGGGCACACCCCGCATGCC | 72 |
| | | | |
| Db | 61 | GCAGCGTCCCGCTGCCGAGCCGTGCCTCCCTGCTGCGCAGACCACTACCGCAGGTGCT | 120 |
| | | | |
| Qy | 73 | CGCGCTCCCGCTGCCGAGCCGTGCCTCCCTGCTGCGCAGACCACTACCGCAGGTGCT | 132 |
| | | | |
| Db | 121 | GCCGCTGCCACGTTCTGTCGGGGCGCTTGGGGCCCCAGGGCTGCGCTGTGTGACGGCGCG | 180 |
| Qy | 133 | GCCGCTGCCACGTTCTGTCGGGGCGCTTGGGGCCCCAGGGCTGCGCTGTGTGACGGCGCG | 192 |
| | | | |
| Db | 181 | GGACCCGGGCGCTTTCGCGCGCTGTGTGGCCCAAGTGCCTGGTGTGGTGCCTTGGACGC | 240 |
| Qy | 193 | GGACCCGGGCGCTTTCGCGCGCTGTGTGGCCCAAGTGCCTGGTGTGGTGCCTTGGACGC | 252 |

1333 AGCGGTGCTGTGCGCGGAGAGCCCGAGGGCTGTGTGGCGCCCCCGAGGAGGGA 1392
1381 CACAGACCCCCGCTGCGCTGTGTGAGCTGCTCCGCCACAGCAGAGCCCTTGGCAGGTGTA 1440
1393 CACAGACCCCCGCTGCGCTGTGTGAGCTGCTCCGCCACAGCAGAGCCCTTGGCAGGTGTA 1452
1441 CGGGTTCGTGGGCGCTGCGCTGCGCGGCTGTGTGCGCCCCAGGCGCTGTGGGGCTCCAGGCA 1500
1453 CGGGTTCGTGGGCGCTGCGCTGCGCGGCTGTGTGCGCCCCAGGCGCTGTGGGGCTCCAGGCA 1512
1501 CAAGAACCGCGCTTCTCAGGAACACCAAGAGTTTCATCTCCTGGGAGCATGCCAA 1560
1513 CAAGAACCGCGCTTCTCAGGAACACCAAGAGTTTCATCTCCTGGGAGCATGCCAA 1572
1561 GCTCTCCTCAGAGAGTACGTGGAGATGAGCGTGGCGGACTGCGCTTGGCTGGCGAG 1620
1573 GCTCTCCTCAGAGAGTACGTGGAGATGAGCGTGGCGGACTGCGCTTGGCTGGCGAG 1632
1621 GAGCCAGAGGTTGGCTGTGTTCGGGCGCAGAGACCGCTCTCGGTGAGAGATCTCTGGC 1680
1633 GAGCCAGAGGTTGGCTGTGTTCGGGCGCAGAGACCGCTCTCGGTGAGAGATCTCTGGC 1692
1681 CAAGTTCTCCTCAGTGTGATGAGTGTACGTGCTGCGAGCTGCTCAGGTCCTTCTTTTA 1740
1693 CAAGTTCTCCTCAGTGTGATGAGTGTACGTGCTGCGAGCTGCTCAGGTCCTTCTTTTA 1752
1741 TGTACAGGAGACCACTTTTCAAAGAACAGCGCTTTTTCACCGAAGAGTGTCTGGAG 1800
1753 TGTACAGGAGACCACTTTTCAAAGAACAGCGCTTTTTCACCGAAGAGTGTCTGGAG 1812
1801 CAAGTTCAAAGCATTTGGAATCAGACAGCACTTTGAGAGGGTGCAGCTGGCGAGTGTCT 1860
1813 CAAGTTCAAAGCATTTGGAATCAGACAGCACTTTGAGAGGGTGCAGCTGGCGAGTGTCT 1872
1861 GGAAGCAGAGTGCAGGAGCATCGGGAAGCCAGGCGCGCCCTGCTGACGTCACAGACTCCG 1920
1873 GGAAGCAGAGTGCAGGAGCATCGGGAAGCCAGGCGCGCCCTGCTGACGTCACAGACTCCG 1932
1921 CTTATCCCCAAGCCTGACGGGCTGCGGCGCATTTGTAACATGGACTACGTCGCTGGGAGC 1980
1933 CTTATCCCCAAGCCTGACGGGCTGCGGCGCATTTGTAACATGGACTACGTCGCTGGGAGC 1992
1981 CAGAACCTTCGAGAGAAAGAGGCGGAGCGTCTCACCTCAGGGTGAAGGACTGTGT 2040
1993 CAGAACCTTCGAGAGAAAGAGGCGGAGCGTCTCACCTCAGGGTGAAGGACTGTGT 2052
2041 CAGGGTCTCAACTACGAGCGGCGGCGCGCCCGCTCCTGGGCGCCTCTGCTGTGG 2100
2053 CAGGGTCTCAACTACGAGCGGCGGCGGCGCCCGCTCCTGGGCGCCTCTGCTGTGG 2112
2101 CTTGGAGGATATCCACAGGCGCTGGCGCACCTTCTGCTGCTGCTGCGGCCCGAGGACCC 2160
2113 CTTGGAGGATATCCACAGGCGCTGGCGCACCTTCTGCTGCTGCTGCGGCCCGAGGACCC 2172
2161 GCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGGTAGACACCATCCCCCA 2220
2173 GCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGGTAGACACCATCCCCCA 2232
2221 GGACAGGCTCACGAGGTATCCCGACATCAACACCCAGCGCTTCAAGAGCCA 2280
2233 GGACAGGCTCACGAGGTATCCCGACATCAACACCCAGCGCTTCAAGAGCCA 2292
2281 TCGGTATCGCGTGTTCAGAAAGCCCGCCATGGCGACGTCGCGAAGCGCTTCAAGAGCCA 2340
2293 TCGGTATCGCGTGTTCAGAAAGCCCGCCATGGCGACGTCGCGAAGCGCTTCAAGAGCCA 2352
2341 CGTCTTACCTTGACAGACCTCCAGCGGTACATCGGACAGTTCGTGGCTCACCTGCAGGA 2400
2353 CGTCTTACCTTGACAGACCTCCAGCGGTACATCGGACAGTTCGTGGCTCACCTGCAGGA 2412
2401 GACAGCGCGCTGAGGATGCGCTGTCATCGAGCAGAGCTCTCCTGATGAGCGCCAG 2460

2413 GACCAGCCCGCTGAGGGATGCGGTGCTATCGAGCAGAGCTCCTCCTGAATGAGGCCAG 2472
2461 CAGTGGCCCTTCGAGCGTCTTCTCAGCTTTCATGTGCGCACCGCGTGCATCAGGGG 2520
2473 CAGTGGCCCTTCGAGCGTCTTCTCAGCTTTCATGTGCGCACCGCGTGCATCAGGGG 2532
2521 CAAGTCTCAGTTCAGTGCAGGGATCCCGAGGGCTCCATCTCTCCACGCTGCTCTG 2580
2533 CAAGTCTCAGTTCAGTGCAGGGATCCCGAGGGCTCCATCTCTCCACGCTGCTCTG 2592
2581 CAGCCTGCTCAGCGGACATGAGAACAGCTGTTTGGGGGATTCGGGGGACGGGCT 2640
2593 CAGCCTGCTCAGCGGACATGAGAACAGCTGTTTGGGGGATTCGGGGGACGGGCT 2652
2641 GCTCCTGCTGTTGGTGGATGATTCTTGTGTTGACACCTCACCTCACCGCCGAAAC 2700
2653 GCTCCTGCTGTTGGTGGATGATTCTTGTGTTGACACCTCACCTCACCGCCGAAAC 2712
2701 CTTCTCAGGACCCCTGCTCCGAGGTGCTCCTGAGTANGCTCGTGGTGAATTCGGGAA 2760
2713 CTTCTCAGGACCCCTGCTCCGAGGTGCTCCTGAGTANGCTCGTGGTGAATTCGGGAA 2772
2761 GACAGTGGTGAATTCCTCTGTAGAACAGCGCCCTGGGTGGACGCTTTTGTTCAGAT 2820
2773 GACAGTGGTGAATTCCTCTGTAGAACAGCGCCCTGGGTGGACGCTTTTGTTCAGAT 2832
2821 GCGGCGCCACGGCCTATTCCCTGCTGCGGCTGCTGCTGGATACCGGACCCCTGGAGT 2880
2833 GCGGCGCCACGGCCTATTCCCTGCTGCGGCTGCTGCTGGATACCGGACCCCTGGAGT 2892
2881 GAGAGCGACTACTCCAGTATGCCCGGACCTCCATCAGAGCGAGTCTCACCTTCAACCG 2940
2893 GAGAGCGACTACTCCAGTATGCCCGGACCTCCATCAGAGCGAGTCTCACCTTCAACCG 2952
2941 CGGCTTCAAGGCTGGGAGAACATCGCTCGCAAACTTTTGGGGTCTTCGCGCTGAAGTG 3000
2953 CGGCTTCAAGGCTGGGAGAACATCGCTCGCAAACTTTTGGGGTCTTCGCGCTGAAGTG 3012
3001 TCACAGCCTGTTTCTGGATTTCAGGTGAACAGCCTCCAGAGCGGTGTCACCAACATCTA 3060
3013 TCACAGCCTGTTTCTGGATTTCAGGTGAACAGCCTCCAGAGCGGTGTCACCAACATCTA 3072
3061 CAAATATCTCTGCTGCGAGCGTACAGGTTTTCAGCATGTGTGCTGCACTCCCAATTCA 3120
3073 CAAATATCTCTGCTGCGAGCGTACAGGTTTTCAGCATGTGTGCTGCACTCCCAATTCA 3132
3121 TCAGCAAGTTTGGAGAACCCACATTTTCTGCGCGTCATCTGACAGCGCTCCCT 3180
3133 TCAGCAAGTTTGGAGAACCCACATTTTCTGCGCGTCATCTGACAGCGCTCCCT 3192
3181 CTGCTACTCCATCTGAAAGCCAAAGCAGGGATGTGCTGGGGGCCAAGGGCGCGC 3240
3193 CTGCTACTCCATCTGAAAGCCAAAGCAGGGATGTGCTGGGGGCCAAGGGCGCGC 3252
3241 CCGGCTCTGCGCCTCCGAGGCGGTGAGTGTGTGCCACCAAGCATTCCTGCTCAAGT 3300
3253 CCGGCTCTGCGCCTCCGAGGCGGTGAGTGTGTGCCACCAAGCATTCCTGCTCAAGT 3312
3301 GACTGCACACCGTGTACCTAGCTGCCACTCCTGGGTCACTCAGGACAGCCAGACGA 3360
3313 GACTGCACACCGTGTACCTAGCTGCCACTCCTGGGTCACTCAGGACAGCCAGACGA 3372
3361 GCTGAGTCCGAAAGCTCCCGGGACGACGCTGACTGCGCTGGAGGCGCAGCAACCCGCG 3420
3373 GCTGAGTCCGAAAGCTCCCGGGACGACGCTGACTGCCCTGGAGGCGCAGCAACCCGCG 3432
3421 ACTGCCCTCAGACTTCAAGACCATCCTTGACTGTATGGCCACCCGCCACAGCCAGGCCGA 3480
3433 ACTGCCCTCAGACTTCAAGACCATCCTTGACTGTATGGCCACCCGCCACAGCCAGGCCGA 3492
3481 GAGCAGACACAGCAGCCTGTACGCGCGCTCTACGTCCTCCAGGGGAGGGGGCGCC 3540
3493 GAGCAGACACAGCAGCCTGTACGCGCGGCTCTACGTCCTCCAGGGGAGGGGGCGCC 3552

QY 912 CCACCTCTTTGGAGGTGGCTCTCTGGCAGCGGCCACTCCACCCATCCGTGGGCGGCC 971
Db 963 AGCACACGCGGGCCCCCATCCACATGCGGGCCACACAGCTCCCTGGGACACGCTTGTG 1022
QY 972 AGCACACGCGGGCCCCCATCCACATGCGGGCCACACAGCTCCCTGGGACACGCTTGTG 1031
Db 1023 CCCGGGTAGCGCGAGACCAAGCACTTCTCTACTCTCTAGGCGGACAAAGAGCAGCTGC 1082
QY 1032 CCCGGGTAGCGCGAGACCAAGCACTTCTCTACTCTCTAGGCGGACAAAGAGCAGCTGC 1091
Db 1083 GGCCTCTCTTCTACTAGCTCTCTAGGCGGACAGCTGACTGGCGCTCGGAGGCTCTGTGG 1142
QY 1092 GGCCTCTCTTCTACTAGCTCTCTAGGCGGACAGCTGACTGGCGCTCGGAGGCTCTGTGG 1151
Db 1143 AGACCATTTTCTGGGTTCAGGCGCTTGATGCCAGGAGTCCCGCAGGTTCGCCCGCC 1202
QY 1152 AGACCATTTTCTGGGTTCAGGCGCTTGATGCCAGGAGTCCCGCAGGTTCGCCCGCC 1211
Db 1203 TGGCCACAGCGTACTGGGAAATGGGCGCCCTGTGTTCTGGAGCTGCTTGGGAACACAGGCG 1262
QY 1212 TGGCCACAGCGTACTGGGAAATGGGCGCCCTGTGTTCTGGAGCTGCTTGGGAACACAGGCG 1271
Db 1263 AGTGCCTCTAGCGGTGCTCTCAAGAGCGCACTGCCGCTCGAGAGCTGCGGTCAACCCAG 1322
QY 1272 AGTGCCTCTAGCGGTGCTCTCAAGAGCGCACTGCCGCTCGAGAGCTGCGGTCAACCCAG 1331
Db 1323 CAGCGGTGTCTGTGCCGGGAGAGGCCCAAGGCTCTGTGGCGGCCCGCCGAGGAGG 1382
QY 1332 CAGCGGTGTCTGTGCCGGGAGAGGCCCAAGGCTCTGTGGCGGCCCGCCGAGGAGG 1391
Db 1383 ACACAGACCCCGCTGCTGTGAGCTGCTCCGCGACGACAGCAGCCCTGGGAGGTGT 1442
QY 1392 ACACAGACCCCGCTGCTGTGAGCTGCTCCGCGACGACAGCAGCCCTGGGAGGTGT 1451
Db 1443 ACGGCTTCTGCGGGCGCTGCTGCGGGCTGCTGCCCGAGGCTCTGGGGCTCCAGGC 1502
QY 1452 ACGGCTTCTGCGGGCGCTGCTGCGGGCTGCTGCCCGAGGCTCTGGGGCTCCAGGC 1511
Db 1503 ACAACGACCGCGCTCTCTAGGAAACACCAAGAGTTTCACTCCCTGGGGAAGCATGCCA 1562
QY 1512 ACAACGACCGCGCTCTCTAGGAAACACCAAGAGTTTCACTCCCTGGGGAAGCATGCCA 1571
Db 1563 AGCTCTGCTCGAGAGCTGACGTGGAGATGAGCGTGGGGGCTGGCTTGGGTGGGCA 1622
QY 1572 AGCTCTGCTCGAGAGCTGACGTGGAGATGAGCGTGGGGGCTGGCTTGGGTGGGCA 1631
Db 1623 GGAGCCAGGGGTGGCTGTGTTCCGGCCGACAGACCGCTGCTGCTGAGGAGATCCTGG 1682
QY 1632 GGAGCCAGGGGTGGCTGTGTTCCGGCCGACAGACCGCTGCTGCTGAGGAGATCCTGG 1691
Db 1683 CCAAGTTCTCTGACTGGGTGATGATGTGTACGTGCTGAGCTGCTCAGGTCTTTCTTTT 1742
QY 1692 CCAAGTTCTCTGACTGGGTGATGATGTGTACGTGCTGAGCTGCTCAGGTCTTTCTTTT 1751
Db 1743 ATGTACGGAGACACGTTTCAAAGAACAGCGCTCTTTTCTACCGAAGAGTGTCTGGGA 1802
QY 1752 ATGTACGGAGACACGTTTCAAAGAACAGCGCTCTTTTCTACCGAAGAGTGTCTGGGA 1811
Db 1803 GCAAGTTCAAAGCATTGAATACAGACAGCACTTGAAGAGGTCAGCTCGGGAGCTGT 1862
QY 1812 GCAAGTTCAAAGCATTGAATACAGACAGCACTTGAAGAGGTCAGCTCGGGAGCTGT 1871
Db 1863 CGGAAGCAGAGGTTCAGGACGATCGGGAAGCCAGGCGCCGCTGCTGACGTTCAGACTCC 1922
QY 1872 CGGAAGCAGAGGTTCAGGACGATCGGGAAGCCAGGCGCCGCTGCTGACGTTCAGACTCC 1931
Db 1923 GCTTCATCCCCAAGCTGACGGGCTGGGCGGATTTGAACATGGATACGCTCGTGGAG 1982
QY 1932 GCTTCATCCCCAAGCTGACGGGCTGGGCGGATTTGAACATGGATACGCTCGTGGAG 1991
Db 1983 CCAGAAGCTTCCGAGAGAAAGAGGCGGAGCGTCTCACTCGAGGTGAAGCACTGT 2042
QY 1992 CCAGAAGCTTCCGAGAGAAAGAGGCGGAGCGTCTCACTCGAGGTGAAGCACTGT 2051

Db 2043 TCAGCGTGTCTCAACTAGAGCGGGCGCGGCCCGGCCCTCTTGGGCGCCTCTGTGTGTG 2102
QY 2052 TCAGCGTGTCTCAACTAGAGCGGGCGCGGCCCGGCCCTCTTGGGCGCCTCTGTGTGTG 2111
Db 2103 GCCTGGAGCATATCCACAGAGGCTTGGGCACTTCTGTGTGTGTGTGTGTGTGTGTGTGT 2162
QY 2112 GCCTGGAGCATATCCACAGGCTTGGGCACTTCTGTGTGTGTGTGTGTGTGTGTGTGT 2171
Db 2163 CGCGCGCTGAGCTGTACTTTTCAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2222
QY 2172 CGCGCGCTGAGCTGTACTTTTCAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2231
Db 2223 AGACAGGCTACGAGGTCATCCGAGCATCATCAAAACCCAGAACACGCTACTGTGTGT 2282
QY 2232 AGACAGGCTACGAGGTCATCCGAGCATCATCAAAACCCAGAACACGCTACTGTGTGT 2291
Db 2283 GTGGGTATGCCGTGGTTCAGAGGCGGCCCATGGGACGCTCCGAGGCGCTTCAAGAGCC 2342
QY 2292 GTGGGTATGCCGTGGTTCAGAGGCGGCCCATGGGACGCTCCGAGGCGCTTCAAGAGCC 2351
Db 2343 ACCTCTCTACTTTGACAGACTCCAGCGTACATGCGACAGTTCTGTGTGTGTGTGTGTGT 2402
QY 2352 ACCTCTCTACTTTGACAGACTCCAGCGTACATGCGACAGTTCTGTGTGTGTGTGTGT 2411
Db 2403 AGACAGCGCGCTGAGGATGCGGTGCTCATCGAGCAGAGCTCTCTCCCTGAATGAGGCCA 2462
QY 2412 AGACAGCGCGCTGAGGATGCGGTGCTCATCGAGCAGAGCTCTCTCCCTGAATGAGGCCA 2471
Db 2463 GCAGTGGCCTTTCGAGCTTCTTCTAGCGTTCATGTGTGTGTGTGTGTGTGTGTGTGTGT 2522
QY 2472 GCAGTGGCCTTTCGAGCTTCTTCTAGCGTTCATGTGTGTGTGTGTGTGTGTGTGTGT 2531
Db 2523 GCAAGTCTACGCTCAGTGCAGGATCCCGAGGCTCCATCTCTCCACGCTGTCT 2582
QY 2532 GCAAGTCTACGCTCAGTGCAGGATCCCGAGGCTCCATCTCTCCACGCTGTCT 2591
Db 2583 GCAGCCTGTCTACGCGACATGGAGAACAAAGCTGTTTGGGGGATTCGGCGGGAGCGGC 2642
QY 2592 GCAGCCTGTCTACGCGACATGGAGAACAAAGCTGTTTGGGGGATTCGGCGGGAGCGGC 2651
Db 2643 TGCTCTGCGCTTGGTGGATGATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2702
QY 2652 TGCTCTGCGCTTGGTGGATGATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2711
Db 2703 CCTTCTCAGAGCCCTGCTCCGAGGTGCTCCTGAGTATGCTGCTGTGTGTGTGTGTGTGT 2762
QY 2712 CCTTCTCAGAGCCCTGCTCCGAGGTGCTCCTGAGTATGCTGCTGTGTGTGTGTGTGT 2771
Db 2763 AGACAGTGTGTAACTTCCCTGTAGAAGACAGGCGCTTGGGTGGCAGCGCTTTTGTTCAGA 2822
QY 2772 AGACAGTGTGTAACTTCCCTGTAGAAGACAGGCGCTTGGGTGGCAGCGCTTTTGTTCAGA 2831
Db 2823 TGGCGGCCACAGGCTTATCCCTGCTGGGCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 2882
QY 2832 TGGCGGCCACAGGCTTATCCCTGCTGGGCGCTGCTGTGTGTGTGTGTGTGTGTGTGT 2891
Db 2883 TGCAAGGCGACTACTCCAGTATGCCGAGCTCCATCAGAGCCAGTCTCACCTTCAACC 2942
QY 2892 TGCAAGGCGACTACTCCAGTATGCCGAGCTCCATCAGAGCCAGTCTCACCTTCAACC 2951
Db 2943 CGCGCTTCAAGGCTGGGAGGAACATGCTCGCAAACTCTTTTGGGGTCTTTCGGGCTGAAGT 3002
QY 2952 CGCGCTTCAAGGCTGGGAGGAACATGCTCGCAAACTCTTTTGGGGTCTTTCGGGCTGAAGT 3011
Db 3003 GTACAGCCTTGTTCGTGGATTTGCAAGTGAACAGCTCCAGAGGCTGTGACCAACATCT 3062
QY 3012 GTACAGCCTTGTTCGTGGATTTGCAAGTGAACAGCTCCAGAGGCTGTGACCAACATCT 3071
Db 3063 ACAAGATCTCTGCTGAGGCGTACAGGTTTTCAGCATGTGTGTGTGTGTGTGTGTGTGTGT 3122
QY 3072 ACAAGATCTCTGCTGAGGCGTACAGGTTTTCAGCATGTGTGTGTGTGTGTGTGTGTGT 3131

Db 361 GCTGCTGACAGGGGCGGGGGGCCCCCGAGGCGCTTACCACACAGGTGCGCAGTA 420
Qy |||||
Db 373 GCTGCTGACAGGGGCGGGGGGCCCCCGAGGCGCTTACCACACAGGTGCGCAGTA 432
Qy |||||
Db 421 CCTGCCAACACAGGTGACCGACGACTCGGGGGAGCGGGGCTGCTCTCGG 480
Qy |||||
Db 433 CCTGCCAACACAGGTGACCGACGACTCGGGGGAGCGGGGCTGCTCTCGG 492
Qy |||||
Db 481 CGCGTGGCGACAGCTGCTGTTTCACTGCTGCGACGCTGCGCGTNTTGTCTGGT 540
Qy |||||
Db 493 CGCGTGGCGACAGCTGCTGTTTCACTGCTGCGACGCTGCGCGCTTGTGCTGGT 552
Qy |||||
Db 541 GNTTCCAGCTGCGCTTACCANGTGTGGGGCGCGCGTGTACACAGCTGCGGCTGCMAC 600
Qy |||||
Db 553 GGTCTCCAGCTGCGCTTACCAGTGTGGGGCGCGCGTGTACACAGCTGCGGCGTGCAC 612
Qy |||||
Db 601 TCAGCGCGCGCGCGCGCGCGCGTANTGGACCGCAANGCTCTGGGAT--CCACAGGCG 658
Qy |||||
Db 613 TCAGCGCGCGCGCGCGCGCGCGTAGTGGACCGCGAAGGCGTGTGGATGCGAACAGGCG 672
Qy |||||
Db 659 CTGGAACCATAGCTGACGGAGCGGGGTCCCCCTGGGC-TGCCAGCGCGGGGTGCGAG 717
Qy |||||
Db 673 CTGGAACCATAGCTGACGGAGCGGGTCCCCCTGGGCCTGCCAGCGCGGGTGCAG 732
Qy |||||
Db 718 GAGCGCGGGGCACTGCGACCGCAAGTCTGCCGTTGCCCAAGAGGCGCCAGCGGTGGCG 777
Qy |||||
Db 733 GAGCGCGGGGCACTGCGACCGCAAGTCTGCCGTTGCCAAGAGGCGCCAGCGGTGGCG 792
Qy |||||
Db 778 TGCCCCCTGAGCGCGAGCGCGCGCTTGGCAGGGGTCTGGCGCCACCGGGCAGGAC 837
Qy |||||
Db 793 TGCCCCCTGAGCGCGAGCGAGCGCGCTTGGCAGGGGTCTGGGGCCACCGGGCAGGAC 852
Qy |||||
Db 838 GCCTGACCGAGTACCGCTGCTGTTCTGTGTGTGTCACCTGCCAGACCGCGCAAGAACG 897
Qy |||||
Db 853 GCGTGGACCGAGTACCGCTGCTGTTCTGTGTGTGTCACCTGCCAGACCGCGCAAGAACG 912
Qy |||||
Db 898 CACTCTTTGGAGGTGGCTCTCTGGCACGGCGCACCTCCACCCATCCGTGGGGCGGCA 957
Qy |||||
Db 913 CACTCTTTGGAGGTGGCTCTCTGGCACGGCGCACCTCCACCCATCCGTGGGGCGGCA 972
Qy |||||
Db 958 GCACACCGGGGCG 1016
Qy |||||
Db 973 GCACACCGGGGCG 1032
Qy |||||
Db 1017 CCGGTGTACGCGCGAGCAAGCACTTCTCTACTCTCAGCGCAACAAGNA-CACTGCGN 1075
Qy |||||
Db 1033 CCGGTGTACGCGCGAGCAAGCACTTCTCTACTCTCAGCGCAAGAGGAGCTGCG 1092
Qy |||||
Db 1076 CCCTCTCTCTACTCAATATATCTGAGCGCGACCGCTGACTGCGGTTGCGGAGTTGCTGG 1135
Qy |||||
Db 1093 GCGCTCTCTCTACTCAAGTCTCTGAGCGCGACCGCTGACTGCGGCTCGGAGGCTCGTGA 1152
Qy |||||
Db 1136 AGACANTCTTCTGCTTCCAGCGCTTGGATGCCAGG-ATTCCCCGACGTTGCCCGCGCT 1194
Qy |||||
Db 1153 GACCATCTTCTGGGTTCCAGCGCGTGGATGCCAGGGACTCCCCGAGGTTGCCCGCGCT 1212
Qy |||||
Db 1195 GCGCCAGCGGNTACTGGCAATTCGCGCGCGCTGTTCTGAGGTGCTTGGGAACCAAGCGCA 1254
Qy |||||
Db 1213 GCGCCAGCGCTACTGGCAATTCGCGCGCGCTGTTCTGAGGTGCTTGGGAACCAAGCGCA 1272
Qy |||||
Db 1255 GTGCGCGCTACGGGTGTTCTTCAAGACGACTGCGCGCTGGAGGTGCGGTCACCGCGAGC 1314
Qy |||||
Db 1273 GTGCGCGCTACGGGTGCTCTCTCAAGACGACTGCGCGCTGGAGGTGCGGTCACCGCGAGC 1332
Qy |||||
Db 1315 AGCGGCTGTCTGCGCGGAGAGCGCGCGCGCTCTGCGCGCGCGCGCGCGCGCGCGAGGAGGA 1374
Qy |||||
Db 1333 AGCGGCTGTCTGCGCGGAGAGCGCGCGCGCTCTGCGCGCGCGCGCGCGCGCGAGGAGGGA 1392
Qy |||||
Db 1375 ACACAGACCGCGCTGCTGTGTGACGTGCTCTCCGCGCGCGCGCGCGCGCGCGCGAGGTGT 1434
Qy |||||
Db 1393 -CACAGACCGCGCTGCTGTGTGACGTGCTCTCCGCGCGCGCGCGCGCGCGCGCGAGGTGT 1451

Db 1435 ACGGCTTCGTGCGGGGCGCTGCTGCGCGCGGCTGTGTGCCCCCAGGCGCTCTGGGGCTCCAGGC 1494
Qy |||||
Db 1452 ACGGCTTCGTGCGGGGCGCTGCTGCGCGCGGCTGTGTGCCCCCAGGCGCTCTGGGGCTCCAGGC 1511
Qy |||||
Db 1495 ACAACGAAAGCGCGCTTCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAGCATGCCA 1554
Qy |||||
Db 1512 ACAACGAAAGCGCGCTTCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAGCATGCCA 1571
Qy |||||
Db 1555 AGCTCTCGCTGCAGGAGCTGACGTGGAAGATGACGCTGCGGAGCTGGGCTTGGCTGGGCA 1614
Qy |||||
Db 1572 AGCTCTCGCTGCAGGAGCTGACGTGGAAGATGACGCTGCGGAGCTGCGCTTGGCTGGGCA 1631
Qy |||||
Db 1615 GGAGCCAGGGGTTGGCTGTGTTCCGCGCGCAGAGCACCGTCTCGGTGAGAGATCCTGG 1674
Qy |||||
Db 1632 GGAGCCAGGGGTTGGCTGTGTTCCGCGCGCAGAGCACCGTCTCGGTGAGAGATCCTGG 1691
Qy |||||
Db 1675 CCAAGTTCCTCGCTGCACTGCTGATGAGTGTACGCTGCGAGCTGCTCAGGTCTTTCCTTT 1734
Qy |||||
Db 1692 CCAAGTTCCTCGCTGCACTGCTGATGAGTGTACGCTGCGAGCTGCTCAGGTCTTTCCTTT 1751
Qy |||||
Db 1735 ATGTCACGGAGACCACTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTTGA 1794
Qy |||||
Db 1752 ATGTCACGGAGACCACTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTTGA 1811
Qy |||||
Db 1795 GCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGTCAGCTGCGGGAGCTGT 1854
Qy |||||
Db 1812 GCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGTCAGCTGCGGGAGCTGT 1871
Qy |||||
Db 1855 CGGAAGCAGAGGTCAGACGATCGGGAAGCCAGGCGCGCGCTGCTCAGCTCCAGACTCC 1914
Qy |||||
Db 1872 CGGAAGCAGAGGTCAGACGATCGGGAAGCCAGGCGCGCGCTGCTCAGCTCCAGACTCC 1931
Qy |||||
Db 1915 GCTTCATCCCCAACCCCTGACGGGCTGCGGCGATTGTGAACATGAGTACGCTCTGCGGAG 1974
Qy |||||
Db 1932 GCTTCATCCCCAACCCCTGACGGGCTGCGGCGATTGTGAACATGAGTACGCTCTGCGGAG 1991
Qy |||||
Db 1975 CCAGAAAGTTCCGCGAGAGAAAGAGGCGCGAGCTCTCACTCGAGGTCGAAGCACTGT 2034
Qy |||||
Db 1992 CCAGAAAGTTCCGCGAGAGAAAGAGGCGCGAGCTCTCACTCGAGGTCGAAGCACTGT 2051
Qy |||||
Db 2035 TCAGCGTGTCACTACGAGGCGCGGCGCGCGCGCGCGCTGCTGGGCGCGCTCTGTGTGG 2094
Qy |||||
Db 2052 TCAGCGTGTCACTACGAGGCGCGGCGCGCGCGCGCGCTGCTGGGCGCGCTCTGTGTGG 2111
Qy |||||
Db 2095 GCGTGGAGCATATCCACAGGCGCTGCGGCGACTTCTGCTGCTGCTGCGGCGCGCGAGCC 2154
Qy |||||
Db 2112 GCGTGGAGCATATCCACAGGCGCTGCGGCGACTTCTGCTGCTGCTGCGGCGCGAGCC 2171
Qy |||||
Db 2155 CGCGCGCTGAGCTGACTTGTCAAGGTGATGTGACGGGCGGTACGACACCACTCCCGC 2214
Qy |||||
Db 2172 CGCGCGCTGAGCTGACTTGTCAAGGTGATGTGACGGGCGGTACGACACCACTCCCGC 2231
Qy |||||
Db 2215 AGGACAGGCTCAGGAGGTCATCCCGAGCATCATCAACCCAGAACACGCTACTGCGTGC 2274
Qy |||||
Db 2232 AGGACAGGCTCAGGAGGTCATCCCGAGCATCATCAACCCAGAACACGCTACTGCGTGC 2291
Qy |||||
Db 2275 GTCGGTATGCGGTGCTCCAGAAAGCGCGCATGGGCACTCCCGCAAGCGCTTCAAGAGCC 2334
Qy |||||
Db 2292 GTCGGTATGCGGTGCTCCAGAAAGCGCGCATGGGCACTCCCGCAAGCGCTTCAAGAGCC 2351
Qy |||||
Db 2335 ACGTCTTACTTTGACAGACTTCCAGCGGTACATGCGACAGTTCGTGGCTACCTGAGG 2394
Qy |||||
Db 2352 ACGTCTTACTTTGACAGACTTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTGAGG 2411
Qy |||||
Db 2395 ANACAGCGCGCTGAGGATGCGTCTATCGAGCAGAGCTCTCTCCCTGNAATGAGGCCA 2454
Qy |||||
Db 2412 AGACAGCGCGCTGAGGATGCGTCTATCGAGCAGAGCTCTCTCCCTGNAATGAGGCCA 2471
Qy |||||
Db 2455 GCAAGTTCCTTTCGAGGCTTCTCTACGCTTCACTGTCACCGCGCGCTGCGCATCAGGG 2514
Qy |||||
Db 2472 GCAAGTTCCTTTCGAGGCTTCTCTACGCTTCACTGTCACCGCGCGCTGCGCATCAGGG 2531
Qy |||||
Db 2515 GCAAGTTCCTTTCGAGGCTTCTCTACGCTTCACTGTCACCGCGCGCTGCGCATCAGGG 2574

| | | | | |
|----|------|-----|---|------|
| QY | 2532 | GC | AAGTCTACGTTCCAGTCCAGGGGATCCCGAGGGCTCCATCCTCTCCACGCTGCTCT | 2591 |
| Db | 2575 | GC | AGCCTGTGCTACGGCGACATCGAGAACAAAGCTGTTTGGCGGGATTTCGGCGGACGCGG | 2634 |
| QY | 2592 | GC | AGCCTGTGCTACGGCGACATGGAGAACAAAGCTGTTTGGCGGGATTTCGGCGGACGCGG | 2651 |
| Db | 2635 | TG | CTCTCTGGGTTTGGTGGATGATTTCTTGTGTGGTACACCTTCACCTCACCCACGCGAAAA | 2694 |
| QY | 2652 | TG | CTCTCTGGGTTTGGTGGATGATTTCTTGTGTGGTACACCTTCACCTCACCCACGCGAAAA | 2711 |
| Db | 2695 | CC | TTCTCAGGACCTGGTCCGAGGTGTCCTGAGTATGAGTGGCTGCGTGAAGTTCGCGA | 2754 |
| QY | 2712 | CC | TTCTCAGGACCTGGTCCGAGGTGTCCTGAGTATGAGTGGCTGCGTGAAGTTCGCGA | 2771 |
| Db | 2755 | AG | CAGTGTGTGAATTTCCCTGTAGAGAGCAGAGCCCTGGGTGGCACAGGCTTTTGTTCAGA | 2814 |
| QY | 2772 | AG | CAGTGTGTGAATTTCCCTGTAGAGAGCAGAGCCCTGGGTGGCACAGGCTTTTGTTCAGA | 2831 |
| Db | 2815 | TG | CGGGCCACAGGCGCTATTCCCTCTGGTGGGCGCTGCTCTGATACCCGAGACCTGGAGG | 2874 |
| QY | 2832 | TG | CGGGCCACAGGCGCTATTCCCTCTGGTGGGCGCTGCTCTGATACCCGAGACCTGGAGG | 2891 |
| Db | 2875 | TG | CAGAGCGACACTCTACGATATGCCGAGACCTTCATCAGAGCCAGTCTCACCTTCAACC | 2934 |
| QY | 2892 | TG | CAGAGCGACACTCTACGATATGCCGAGACCTTCATCAGAGCCAGTCTCACCTTCAACC | 2951 |
| Db | 2935 | GC | GGCTTTCAAGGCTGGAGGAACATGCGTCGCGAAACTCTTTGGGGTCTTTGGCGGTGAAGT | 2994 |
| QY | 2952 | GC | GGCTTTCAAGGCTGGAGGAACATGCGTCGCGAAACTCTTTGGGGTCTTTGGCGGTGAAGT | 3011 |
| Db | 2995 | GT | CACAGCCTGTTTCTGSAATTTGCAGGTGAACAGAGCTCCAGACAGGTGTGCACCAACATCT | 3054 |
| QY | 3012 | GT | CACAGCCTGTTTCTGSAATTTGCAGGTGAACAGAGCTCCAGACAGGTGTGCACCAACATCT | 3071 |
| Db | 3055 | ACA | GATCCTCTGCTGCAGGGGTACAGGTTTCAGCGATGTGCTGTGAGCTCCCATTTTC | 3114 |
| QY | 3072 | ACA | GATCCTCTGCTGCAGGGGTACAGGTTTCAGCGATGTGCTGTGAGCTCCCATTTTC | 3131 |
| Db | 3115 | AT | CAGAAAGTTTGGAAAGAACCCACATTTTCTCGCGGTCTATCTCTGACAGGCGCTCC | 3174 |
| QY | 3132 | AT | CAGAAAGTTTGGAAAGAACCCACATTTTCTCGCGGTCTATCTCTGACAGGCGCTCC | 3191 |
| Db | 3175 | TC | TGCTACTCCTATCTTCAAGCCAAAGACGACAGGATGTCTGTGGGGCCCAAGGGCGCG | 3234 |
| QY | 3192 | TC | TGCTACTCCTATCTTCAAGCCAAAGACGACAGGATGTCTGTGGGGCCCAAGGGCGCG | 3251 |
| Db | 3235 | CG | GGCCCTCTGGCCCTCCGAGGCGGTGCAGTGGCTGTGCCACCAAGCATTTCTGTCTAAGC | 3294 |
| QY | 3252 | CG | GGCCCTCTGGCCCTCCGAGGCGGTGCAGTGGCTGTGCCACCAAGCATTTCTGTCTAAGC | 3311 |
| Db | 3295 | TG | ACTCGACACCGTGTACCTAGTGTCCACTCTGGGGTCACTCAGACAGCCAGACGC | 3354 |
| QY | 3312 | TG | ACTCGACACCGTGTACCTAGTGTCCACTCTGGGGTCACTCAGACAGCCAGACGC | 3371 |
| Db | 3355 | AG | TGAGTCGAAAGCTCCCGGGAGACGCTGACTGTCCCTGGAGCGCGACGCAACCCGG | 3414 |
| QY | 3372 | AG | TGAGTCGAAAGCTCCCGGGAGACGCTGACTGTCCCTGGAGCGCGACGCAACCCGG | 3431 |
| Db | 3415 | CA | TGCGCCTCAGACTTCAAGACCATCTTGAGTGTATGGCCACCCGCCACAGCCAGGCGG | 3474 |
| QY | 3432 | CA | TGCGCCTCAGACTTCAAGACCATCTTGAGTGTATGGCCACCCGCCACAGCCAGGCGG | 3491 |
| Db | 3475 | AG | ACAGACACACAGACGCCCTGTCAGCGCGGCTCTAGTCTCAGAGGAGGAGGGGGCGG | 3534 |
| QY | 3492 | AG | ACAGACACACAGACGCCCTGTCAGCGCGGCTCTAGTCTCAGAGGAGGAGGGGGCGG | 3551 |
| Db | 3535 | CC | ACACCGGCGCCGACCGCTGGAGTCTGAGGCGCTGAGTGAAGTGTGTGGCCGAGGCGCT | 3594 |
| QY | 3552 | CC | ACACCGGCGCCGACCGCTGGAGTCTGAGGCGCTGAGTGAAGTGTGTGGCCGAGGCGCT | 3611 |
| Db | 3595 | GC | ATGTCGCGCTGAAGGCTGAGTGTGCCGCTGAGGCGCTGAGCGAGTGTCCAGCAAGGCG | 3654 |

| | | | |
|--------|--|--|------|
| Qy | 3612 | CGATGTCGGGCTGAAGGCTGAGTGTCCGGCTGAGGCGCTGAGGGAGTGTCCACCAAGGGC | 3671 |
| Db | 3655 | TGAGTGTCCAGCACACCTGCCGCTTTCACCTCCCCACAGGCTGGCGCTCGGCTCCACCCC | 3714 |
| Qy | 3672 | TGAGTGTCCAGCACACCTGCCGCTTTCACCTCCCCACAGGCTGGCGCTCGGCTCCACCCC | 3731 |
| Db | 3715 | AGGCCAGCTTTTCTTCACAGAGAGCCGGCTTCCACTCCCCACATAGGAATAGTCCATC | 3774 |
| Qy | 3732 | AGGCCAGCTTTTCTTCACAGAGAGCCGGCTTCCACTCCCCACATAGGAATAGTCCATC | 3791 |
| Db | 3775 | CCAGAT 3781 | |
| Qy | 3792 | CCCTGAT 3798 | |
| RESULT | 8 | | |
| ID | US-08-851-843-173 | STANDARD; DNA; UNC; 4029 BP. | |
| AC | XXXXX | | |
| DE | Sequence 173, Application US/08851843 | | |
| CC | Sequence 173, Application US/08851843 | | |
| CC | GENERAL INFORMATION: | | |
| CC | APPLICANT: Cech, Thomas R. | | |
| CC | APPLICANT: Lingner, Joachim | | |
| CC | APPLICANT: Nakamura, Toru | | |
| CC | APPLICANT: Chapman, Karen B. | | |
| CC | APPLICANT: Morin, Gregg B. | | |
| CC | APPLICANT: Harley, Calvin | | |
| CC | APPLICANT: Andrews, William H. | | |
| CC | TITLE OF INVENTION: Novel Telomerase | | |
| CC | NUMBER OF SEQUENCES: 223 | | |
| CC | CORRESPONDENCE ADDRESS: | | |
| CC | ADDRESSEE: Townsend and Townsend and Crew LLP | | |
| CC | STREET: Two Embarcadero Center, 8th Floor | | |
| CC | CITY: San Francisco | | |
| CC | STATE: California | | |
| CC | COUNTRY: United States of America | | |
| CC | ZIP: 94111 | | |
| CC | COMPUTER READABLE FORM: | | |
| CC | MEDIUM TYPE: Floppy disk | | |
| CC | COMPUTER: IBM PC compatible | | |
| CC | OPERATING SYSTEM: PC-DOS/MS-DOS | | |
| CC | SOFTWARE: PatentIn Release #1.0, Version #1.30 | | |
| CC | CURRENT APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: US/08/851,843 | | |
| CC | FILING DATE: 06-MAY-1997 | | |
| CC | CLASSIFICATION: 536 | | |
| CC | PRIOR APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: US 08/846,017 | | |
| CC | FILING DATE: 25-APR-1997 | | |
| CC | CLASSIFICATION: 536 | | |
| CC | PRIOR APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: US 08/844,419 | | |
| CC | FILING DATE: 18-APR-1997 | | |
| CC | CLASSIFICATION: 536 | | |
| CC | PRIOR APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: US 08/724,643 | | |
| CC | FILING DATE: 01-OCT-1996 | | |
| CC | CLASSIFICATION: 536 | | |
| CC | ATTORNEY/AGENT INFORMATION: | | |
| CC | NAME: Apple, Randolph T. | | |
| CC | REGISTRATION NUMBER: 36,429 | | |
| CC | REFERENCE/DOCKET NUMBER: 015389-00293005 | | |
| CC | TELECOMMUNICATION INFORMATION: | | |
| CC | TELEPHONE: (415) 576-0200 | | |
| CC | TELEFAX: (415) 576-0300 | | |
| CC | INFORMATION FOR SEQ ID NO: 173: | | |
| CC | SEQUENCE CHARACTERISTICS: | | |
| CC | LENGTH: 4029 base pairs | | |
| CC | TYPE: nucleic acid | | |
| CC | STRANDEDNESS: single | | |
| CC | TOPOLOGY: linear | | |
| CC | MOLECULE TYPE: CDNA | | |

QY 1992 CCAGAACGTTCCGCGACAGAAAAGAGGGCGGAGCGTCTCACCTCGAGGGTGAAGCACTGT 2051
Db 2035 TACAGGTGCTCAACTACAGCGGGCGCGCGCCCGGCTCTCTGGCGCCTCTGTGCTGG 2094
QY 2052 TCAGCGTCTCAACTACGAGCGGGCGCGCGCCCGGCTCTCTGGCGCCTCTGTGCTGG 2111
Db 2095 GCCTGGAGATATCCACAGGGCCTGGCGCACCTTCTGCTGCTGGGTGTGGGGCCAGGACC 2154
QY 2112 GCCTGGAGATATCCACAGGGCCTGGCGCACCTTCTGCTGCTGGGTGTGGGGCCAGGACC 2171
Db 2155 GCGCCCTGAGCTGTACTTTGTCAGGTGGATGTACGGGGCGGTACGACACCATCCCCC 2214
QY 2172 GCGCCCTGAGCTGTACTTTGTCAGGTGGATGTACGGGGCGGTACGACACCATCCCCC 2231
Db 2215 AGGACAGGCTCACGGAGTCTACGCCAGCATCATCAAAACCCAGAACACGCTACGCTGC 2274
QY 2232 AGGACAGGCTCACGGAGTCTACGCCAGCATCATCAAAACCCAGAACACGCTACGCTGC 2291
Db 2275 GTCGATATCCGCTGTCTCAGAAAGCCGCCCATGGCGACGTCCGGAAGGCTTCAAGAGCC 2334
QY 2292 GTCGATATCCGCTGTCTCAGAAAGCCGCCCATGGCGACGTCCGGAAGGCTTCAAGAGCC 2351
Db 2335 AGCTCTCTACCTTTGACAGACCTCCAGCCGTACATCGGACAGTTCTGTGCTACCTGCGAGG 2394
QY 2352 AGCTCTCTACCTTTGACAGACCTCCAGCCGTACATCGGACAGTTCTGTGCTACCTGCGAGG 2411
Db 2395 ANAACAGCCCGCTGAGGATCCGCTCATCGAGCAGAGTCTCCCTGATGAGGCCA 2454
QY 2412 AGAACAGCCCGCTGAGGATCCGCTCATCGAGCAGAGTCTCCCTGATGAGGCCA 2471
Db 2455 GCAGTGGCCTCTTCAGCTCTTCTTACGCTTTCATGTGTCACACACCGCTGGCATCAGGG 2514
QY 2472 GCAGTGGCCTCTTTCAGCTCTTCTTACGCTTTCATGTGTCACACACCGCTGGCATCAGGG 2531
Db 2515 GCAAGTCTCTACGTCAGTCCAGAGGAGTCCCGCAGGGCTCCATCTCTCCAGGCTGCTCT 2574
QY 2532 GCAAGTCTCTACGTCAGTCCAGAGGAGTCCCGCAGGGCTCCATCTCTCCAGGCTGCTCT 2591
Db 2575 GCAGCTGTGCTACGCGACATGAGAACAAAGCTGTTTGGGGGATTCGGGGGACGGGC 2634
QY 2592 GCAGCTGTGCTACGCGACATGAGAACAAAGCTGTTTGGGGGATTCGGGGGACGGGC 2651
Db 2635 TGCTCTGGTTTGGTGAATTTCTTGTGGTGACACCTCACCTCACCCACGCGAA 2694
QY 2652 TGCTCTGGTTTGGTGAATTTCTTGTGGTGACACCTCACCTCACCCACGCGAA 2711
Db 2695 CTTCTCTCAGGACCTGTGTCAGAGTCCCTGAGTATGGCTGCTGAGTTCGCGA 2754
QY 2712 CTTCTCTCAGGACCTGTGTCAGAGTTCCTTGGTATGGCTGCTGAGTTCGCGA 2771
Db 2755 AGACAGTGTGAATTCCTGTAGAGACGAGCGCCTGGTGGCACGGCTTTTGTTCAGA 2814
QY 2772 AGACAGTGTGAATTCCTGTAGAGACGAGCGCCTGGTGGCACGGCTTTTGTTCAGA 2831
Db 2815 TGCCGGCCACGGCTATTCCTCTGGTGGGGCTCTGCTGATACCCGGACCTGGAGG 2874
QY 2832 TGCCGGCCACGGCTATTCCTCTGGTGGGGCTCTGCTGATACCCGGACCTGGAGG 2891
Db 2875 TGACAGCGACTACTCCAGTATCCCGACCTCCATCAGACCCAGTCTCACCTTCAACC 2934
QY 2892 TGACAGCGACTACTCCAGTATCCCGACCTCCATCAGACCCAGTCTCACCTTCAACC 2951
Db 2935 GCGGCTTCAAGGCTGGGAGAACATCGCTGCAAACTCTTTGGGCTCTTGGGCTGAAGT 2994
QY 2952 GCGGCTTCAAGGCTGGGAGAACATCGCTGCAAACTCTTTGGGCTCTTGGGCTGAAGT 3011
Db 2995 GTACAGCGCTGTTCTGATTTGAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCT 3054
QY 3012 GTACAGCGCTGTTCTGATTTGAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCT 3071
Db 3055 ACAAGATCTCTGCTGAGGGGTACAGTTTACGCGATGTGTGCTGAGCTCCCATTC 3114
QY 3072 ACAAGATCTCTGCTGAGGGGTACAGTTTACGCGATGTGTGCTGAGCTCCCATTC 3131

Db 3115 ATCAGCAAGTTTGAAGAACCACATTTTCTGCGGTCTCTGTGACAGCGCTCC 3174
QY 3132 ATCAGCAAGTTTGAAGAACCACATTTTCTGCGGTCTCTGTGACAGCGCTCC 3191
Db 3175 TCTGCTACTCCTCTGAAAGCCAAAGAACGAGGATGCTGCTGGGGCCAAAGGGCGCG 3234
QY 3192 TCTGCTACTCCTCTGAAAGCCAAAGAACGAGGATGCTGCTGGGGCCAAAGGGCGCG 3251
Db 3235 CCGGCGCTCTCCCTCCGAGGCGGTGAGTGGCTGTGCACCAAGCATTTCTGTCTCAAGC 3294
QY 3252 CCGGCGCTCTCCCTCCGAGGCGGTGAGTGGCTGTGCACCAAGCATTTCTGTCTCAAGC 3311
Db 3295 TGACTCGACACCGTGTACCTTACCTGACCTGCTGGGTCTACTCAGGACAGCCAGAGCG 3354
QY 3312 TGACTCGACACCGTGTACCTTACCTGACCTGCTGGGTCTACTCAGGACAGCCAGAGCG 3371
Db 3355 AGCTGAGTCGGAAGCTCCCGGGGAGGAGCTGACTGCCCTGGAGCGCCAGCAACCGCG 3414
QY 3372 AGCTGAGTCGGAAGCTCCCGGGGAGGAGCTGACTGCCCTGGAGCGCCAGCAACCGCG 3431
Db 3415 CACTGCCCTCAGACTTCAAGACCATCTCTGGACTGATGGCCACCGCCACAGCCAGAGCG 3474
QY 3432 CACTGCCCTCAGACTTCAAGACCATCTCTGGACTGATGGCCACCGCCACAGCCAGAGCG 3491
Db 3475 AGAGCAGACACCGAGCCCTGTCAAGCCGGCTTACGTCAGGGAGGAGGGCGCG 3534
QY 3492 AGAGCAGACACCGAGCCCTGTCAAGCCGGCTTACGTCAGGGAGGAGGGCGCG 3551
Db 3535 CCACACCGAGCCCGCAGCCCTGGAGTCTGAGGCTGAGTGTGTTGGCCAGGCT 3594
QY 3552 CCACACCGAGCCCGCAGCCCTGGAGTCTGAGGCTGAGTGTGTTGGCCAGGCT 3611
Db 3595 GCATGTCGGCTGAAGGCTGAGTGTGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3654
QY 3612 GCATGTCGGCTGAAGGCTGAGTGTGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3671
Db 3655 TGAGTGTCCAGCACACCTTCCCTTCACTTCCACAGGCTGCGCTGCGCTCCACGCC 3714
QY 3672 TGAGTGTCCAGCACACCTTCCCTTCACTTCCACAGGCTGCGCTGCGCTCCACGCC 3731
Db 3715 AGGGCAGCTTTCTTCCACAGGAGCCGGCTTCCACTTCCACATAGATAGTTCATC 3774
QY 3732 AGGGCAGCTTTCTTCCACAGGAGCCGGCTTCCACTTCCACATAGATAGTTCATC 3791
Db 3775 CCCAGAT 3781
QY 3792 CCTGAT 3798

RESULT 9

ID US-08-912-951-4 STANDARD; DNA; UNC; 3855 BP.

AC xxxxx

DE Sequence 4, Application US/08912951

CC Sequence 4, Application US/08912951

CC GENERAL INFORMATION:

CC APPLICANT: Cech, Thomas R.

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin

CC APPLICANT: Andrews, William H.

CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND

CC NUMBER OF SEQUENCES: 335

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Townsend and Crew LLP

CC STREET: Two Embarcadero Center, 8th Floor

CC CITY: San Francisco

CC STATE: California

CC COUNTRY: United States of America

|||||
QY 1273 GTGCCCCACGGGTGCTCTCAAGACGCACCTGCCCGCTGGAGCTGGGTCACCCCAGC 1332
Db |||||
QY 1321 AGCCGGTGTCTGTGCCCGGAGAGCCCGAGGGCTCTGTGGCGGCCCGCCGAGGAGGAGGA 1380
Db |||||
QY 1333 AGCCGGTGTCTGTGCCCGGAGAGCCCGAGGGCTCTGTGGCGGCCCGCCGAGGAGGAGGA 1392
Db |||||
QY 1381 CACAGACCCCGTGGCTGTGTGAGCTGTCTCGGCCACACAGAGCCCGTGGCAGGTGTA 1440
Db |||||
QY 1393 CACAGACCCCGTGGCTGTGTGAGCTGTCTCGGCCACACAGAGCCCGTGGCAGGTGTA 1452
Db |||||
QY 1441 CGGCTTGTGGGCGCTGTCTGCGCGGGCTGTGCGCCCGAGGCTCTGGGCTCCAGGCA 1500
Db |||||
QY 1453 CGGCTTGTGGGCGCTGTCTGCGCGGGCTGTGCGCCCGAGGCTCTGGGCTCCAGGCA 1512
Db |||||
QY 1501 CAAGCAACCGCGCTTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAA 1560
Db |||||
QY 1513 CAAGCAACCGCGCTTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAA 1572
Db |||||
QY 1561 GCTCTCGCTGAGAGCTGAGCTGGAAGATGAGCGTGGGGAGTGCCTTGGCTGCGCAG 1620
Db |||||
QY 1573 GCTCTCGCTGAGAGCTGAGCTGGAAGATGAGCGTGGGGAGTGCCTTGGCTGCGCAG 1632
Db |||||
QY 1621 GAGCCAGGGTGTGGCTGTGTTCCGCGCGAGAGCACCGCTGCTGAGGAGATCCTGGC 1680
Db |||||
QY 1633 GAGCCAGGGTGTGGCTGTGTTCCGCGCGAGAGCACCGCTGCTGAGGAGATCCTGGC 1692
Db |||||
QY 1681 CAAGTTCTGCACCTGGCTGATGAGTGTACGTCGTCGAGCTGCTCAGGCTTTCCTTTTA 1740
Db |||||
QY 1693 CAAGTTCTGCACCTGGCTGATGAGTGTACGTCGTCGAGCTGCTCAGGCTTTCCTTTTA 1752
Db |||||
QY 1741 TGTCAAGGAGACACGTTTCAAAAGAACAGGCTCTTTTTCACCGGAAGAGTGTCTGGAG 1800
Db |||||
QY 1753 TGTCAAGGAGACACGTTTCAAAAGAACAGGCTCTTTTTCACCGGAAGAGTGTCTGGAG 1812
Db |||||
QY 1801 CAAGTTCAAGCATTTGGAATCACAGACACTTGAAGAGGCTGACGCTGCGGGAGCTGTC 1860
Db |||||
QY 1813 CAAGTTCAAGCATTTGGAATCACAGACACTTGAAGAGGCTGACGCTGCGGGAGCTGTC 1872
Db |||||
QY 1861 GGAAGCAGAGTCAAGGAGCATCGGGAAGCCAGGCGCCCTGCTGACGCTCCAGACTCCG 1920
Db |||||
QY 1873 GGAAGCAGAGTCAAGGAGCATCGGGAAGCCAGGCGCCCTGCTGACGCTCCAGACTCCG 1932
Db |||||
QY 1921 CTTCATCCCCAAGCCTGACGGGCTGCGGCGGATTTGTAACATGAGACTACGCTGCGGAGC 1980
Db |||||
QY 1933 CTTCATCCCCAAGCCTGACGGGCTGCGGCGGATTTGTAACATGAGACTACGCTGCGGAGC 1992
Db |||||
QY 1981 CAGAACGTTCCGCAGAGAAAGAGGCGCGGAGCTCTACCTCGAGGCTGAAGGACACTGTT 2040
Db |||||
QY 1993 CAGAACGTTCCGCAGAGAAAGAGGCGCGGAGCTCTACCTCGAGGCTGAAGGACACTGTT 2052
Db |||||
QY 2041 CAGGCTGCTCAACTACGAGCGGCGGCGCGCCGCTCTGCGGCGCTCTGCTGCTGGG 2100
Db |||||
QY 2053 CAGGCTGCTCAACTACGAGCGGCGGCGCGCCGCTCTGCGGCGCTCTGCTGCTGGG 2112
Db |||||
QY 2101 CCTGAGCATATCACAGGGCTGCGGCGACCTCTGCTGCTGCTGCGGCGCCAGGACCC 2160
Db |||||
QY 2113 CCTGAGCATATCACAGGGCTGCGGCGACCTCTGCTGCTGCTGCGGCGCCAGGACCC 2172
Db |||||
QY 2161 GCCGCTGAGCTGACTTTGTCAAGGTGGATGTGACGGGCGGCTACGACACCATCCGCCA 2220
Db |||||
QY 2173 GCCGCTGAGCTGACTTTGTCAAGGTGGATGTGACGGGCGGCTACGACACCATCCGCCA 2232
Db |||||
QY 2221 GGACAGGCTCAGGAGGTCAATCGCAGCATCATCAAAACCCAGAACACTACTGCGTGGC 2280
Db |||||
QY 2233 GGACAGGCTCAGGAGGTCAATCGCAGCATCATCAAAACCCAGAACACTACTGCGTGGC 2292
Db |||||
QY 2281 TCGGTATGCCGTGTCCAGAGGCGCGCCATGGCAGCTCCGCAAGGCTTCAAGAGCCA 2340
Db |||||
QY 2293 TCGGTATGCCGTGTCCAGAGGCGCGCCATGGCAGCTCCGCAAGGCTTCAAGAGCCA 2352
Db |||||
QY 2341 CGTC----- 2344
Db |||||

QY 2353 CGTCTCTACCTTGACAGACCTCCAGCGGTACATGCGACAGTTCTGTTGGCTACCTGCGAGGA 2412
Db |||||
QY 2413 GACCAGCCCGCTGAGGGATGCCGTGCTCATGAGCAGAGCTCCTCCTGAATGAGGCCAG 2472
Db |||||
QY 2473 CAGTGGCTTCTGACGCTTCTCTACGCTTCATGTGCCACACGCGGTGCGCATCAGGGG 2532
Db |||||
QY 2533 CAGTCTCTGCTGAGGGATGCCGTGCTCATGAGCAGAGCTCCTCCTGAATGAGGCCAG 2592
Db |||||
QY 2593 CAGCCTGTGTACGGGACATGGAGAACAAAGCTGTTTGGGGGATTCGGCGGACGGCT 2652
Db |||||
QY 2653 GCTCTGCTGCTGAGGGATGATTTCTTGTGTTGGTACACCTCAGCTCACCACCGGAAAC 2712
Db |||||
QY 2713 CTTCTCAGAGACCTGCTCGAGGTGCTCCTGAGTATGGCTGCTGAGTGAACTTGCGGAA 2772
Db |||||
QY 2773 GACAGTGTGAACTTCCCTGTAGAAGACGAGGCCCTGGGTGGACAGGCTTTTGTTCAGAT 2832
Db |||||
QY 2833 GCGGGCCACGGCTATTCCTCTGAGAGAGAGAGGCCCTGGGTGGACAGGCTTTTGTTCAGAT 2892
Db |||||
QY 2893 GCGGGCCACGGCTATTCCTCTGAGAGAGAGAGGCCCTGGGTGGACAGGCTTTTGTTCAGAT 2952
Db |||||
QY 2953 GCGGGCCACGGCTATTCCTCTGAGAGAGAGAGGCCCTGGGTGGACAGGCTTTTGTTCAGAT 3012
Db |||||
QY 2993 GCGGGCCACGGCTATTCCTCTGAGAGAGAGAGGCCCTGGGTGGACAGGCTTTTGTTCAGAT 3072
Db |||||
QY 3073 CAAGATCCTCTGCTGAGGGCTACAGGTTTCAGCGCATGTGTGCTGCGAGCTTCCCATTTCA 3132
Db |||||
QY 3133 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGGCTCATCTCTGACACGCGCTCCCT 2998
Db |||||
QY 3193 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGGCTCATCTCTGACAGGCTCCCT 3192
Db |||||
QY 3293 CTGCTACTCCATCCTGAAAGCCCAAGAACGCGAGGATGCTGCTGGGGCCCAAGGCGCGCG 3058
Db |||||
QY 3393 CTGCTACTCCATCCTGAAAGCCCAAGAACGCGAGGATGCTGCTGGGGCCCAAGGCGCGCG 3252
Db |||||
QY 3493 CGGCGCTCTGCCCTCCGAGGCGGTGAGTGGCTGCGCACCAAGCATTCCTGCTCAAGCT 3318
Db |||||
QY 3593 CGGCGCTCTGCCCTCCGAGGCGGTGAGTGGCTGCGCACCAAGCATTCCTGCTCAAGCT 3372
Db |||||
QY 3693 GACTTCGACACCGTGTACCTACGTCGCGGCTCATCTGAGGAGAGAGAGAGAGAGCA 3178
Db |||||
QY 3793 GACTTCGACACCGTGTACCTACGTCGCGGCTCATCTGAGGAGAGAGAGAGAGAGCA 3372
Db |||||
QY 3893 GCTGAGTGGAGGCTCCCGGGGAGAGCGTGTGCTGAGTGGAGGCGGAGGCGGAGAGAGCA 3238
Db |||||
QY 3993 GCTGAGTGGAGGCTCCCGGGGAGAGCGTGTGCTGAGTGGAGGCGGAGGCGGAGAGAGCA 3432
Db |||||
QY 4093 ACTGCCCTCAGACTTCAAGACCATCTCTGGAGTGTGAGTGGAGGCGGAGGCGGAGAGAGCA 3298
Db |||||
QY 4193 ACTGCCCTCAGACTTCAAGACCATCTCTGGAGTGTGAGTGGAGGCGGAGGCGGAGAGAGCA 3492

QY 613 TCAGGCCGCGCCGACACACGCTAGTGGACCCCGAAGCGCTCTGGGATGCGAACGGCG 672
Db 661 CTGGAACCATAGGCTCAGGAGCGCGGGTCCCTCTGGGCTTCCAGACCCCGGGTGGAG 720
QY 673 CTGGAACCATAGGCTCAGGAGCGCGGGTCCCTCTGGGCTTCCAGACCCCGGGTGGAG 732
Db 721 GAGCGCGGGGAGTCCAGCGAAGCTCTGCGCTTGCCTCAAGAGGCGCCAGGCTGSCGC 780
QY 733 GAGCGCGGGGAGTCCAGCGAAGCTCTGCGCTTGCCTCAAGAGGCGCCAGGCTGSCGC 792
Db 781 TGCCCTTAGCCGAGCGGAGCGCCGCTTGGGAGGGTCTTGGGCCACCCCGGGCGAGAC 840
QY 793 TGCCCTTAGCCGAGCGGAGCGCCGCTTGGGAGGGTCTTGGGCCACCCCGGGCGAGAC 852
Db 841 GCGTGACCGAGTACCGTGGTCTCTGTGTGTGTACCTGCGCAGACCCCGCGAAGAC 900
QY 853 GCGTGACCGAGTACCGTGGTCTCTGTGTGTGTACCTGCGCAGACCCCGCGAAGAC 912
Db 901 CACCTCTTTGAGGGTSCGCTCTGTGACGCGCCACTCCACCCATCCGCTGGGCGGCCA 960
QY 913 CACCTCTTTGAGGGTSCGCTCTGTGACGCGCCACTCCACCCATCCGCTGGGCGGCCA 972
Db 961 GCACACGCGGGGCCCCCATCCACATCGCGGCCACACGCTCCCTGGACACGCTTGTCC 1020
QY 973 GCACACGCGGGGCCCCCATCCACATCGCGGCCACACGCTCCCTGGACACGCTTGTCC 1032
Db 1021 CCGGGTAGCCGAGACCAAGACATTCCTCTACTCTCAGGGGACCAAGAGCAGTGGC 1080
QY 1033 CCGGGTAGCCGAGACCAAGACATTCCTCTACTCTCAGGGGACCAAGAGCAGTGGC 1092
Db 1081 GCCTCTCTCTACTAGCTCTCTGAGGCCAGCGTCTGAGGCTCGGAGGCTCTGTGA 1140
QY 1093 GCCTCTCTCTACTAGCTCTCTGAGGCCAGCGTCTGAGGCTCGGAGGCTCTGTGA 1152
Db 1141 GACCATCTTTCTGGTTCAGGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCCCT 1200
QY 1153 GACCATCTTTCTGGTTCAGGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCCCT 1212
Db 1201 GCGCCAGCGTACTGGCAATAGGCGCCCTGTCTTCTGGAGCTGTGGGAAACACGCGCA 1260
QY 1213 GCGCCAGCGTACTGGCAATAGGCGCCCTGTCTTCTGGAGCTGTGGGAAACACGCGCA 1272
Db 1261 GTGCCCTACGGGTGCTCTCAGAGCGACTGCCCGCTCGGAGCTCGGCTACCCCGAGC 1320
QY 1273 GTGCCCTACGGGTGCTCTCAGAGCGACTGCCCGCTCGGAGCTCGGCTACCCCGAGC 1332
Db 1321 AGCGGTGTCTGTGCCGGGAGAGCCCGAGGGCTCTGTGGCGGCCCGCAGGAGGAGGA 1380
QY 1333 AGCGGTGTCTGTGCCGGGAGAGCCCGAGGGCTCTGTGGCGGCCCGCAGGAGGAGGA 1392
Db 1381 CACAGACCCCGCTGCTGT 1440
QY 1393 CACAGACCCCGCTGCTGT 1452
Db 1441 CGGCTTGTGGGGCTTGCCTGGCGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
QY 1453 CGGCTTGTGGGGCTTGCCTGGCGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1512
Db 1501 CAACGAACCGCGCTTCTCAGGAACACCAAGAGTTCTCTCCTTGGGGAAGCATGCCAA 1560
QY 1513 CAACGAACCGCGCTTCTCAGGAACACCAAGAGTTCTCTCCTTGGGGAAGCATGCCAA 1572
Db 1561 GCTCTCCTCAGAGAGTACGTGGAGATGAGCGTGGGGAGTCTGCGGACTTGGCTTGGCAG 1620
QY 1573 GCTCTCCTCAGAGAGTACGTGGAGATGAGCGTGGGGAGTCTGCGGACTTGGCTTGGCAG 1632
Db 1621 GAGCCAGGGTGGCTGT 1680
QY 1633 GAGCCAGGGTGGCTGT 1692
Db 1681 CAACTTCTCCTAGCTGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1740
QY 1693 CAACTTCTCCTAGCTGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1752

Db 1741 TGTACGGAGACACACGCTTTCAAAGAACAGGCTCTTTTCTTACCAGAACAGTCTCTGGAG 1800
QY 1753 TGTACGGAGACACACGCTTTCAAAGAACAGGCTCTTTTCTTACCAGAACAGTCTCTGGAG 1812
Db 1801 CAACTTCAAAGCATTTGGAATCAGACAGACCTTTGAAGAGGGTGCAGCTGCGGAGCTGTC 1860
QY 1813 CAACTTCAAAGCATTTGGAATCAGACAGACCTTTGAAGAGGGTGCAGCTGCGGAGCTGTC 1872
Db 1861 GGAAGCAGAGGTTCAGCAGCATCGGGAAGCCAGGCGCCGCTCTGCTGACGTCCAGACTCCG 1920
QY 1873 GGAAGCAGAGGTTCAGCAGCATCGGGAAGCCAGGCGCCGCTCTGCTGACGTCCAGACTCCG 1932
Db 1921 CTTTATCCCAAGCCCTGAGGGCTGCGGGGATTTGTAACATGGACTACGCTCTGGGAGC 1980
QY 1933 CTTTATCCCAAGCCCTGAGGGCTGCGGGGATTTGTAACATGGACTACGCTCTGGGAGC 1992
Db 1981 CAGAAGCTTCCCGCAGAGAAAGAGGCGCCGCTCTGCTGAGGGTGAAGCACTGTT 2040
QY 1993 CAGAAGCTTCCCGCAGAGAAAGAGGCGCCGCTCTGCTGAGGGTGAAGCACTGTT 2052
Db 2041 CAGCGTGTCAACTACGAGCGCGCGCGCCGCTCTGCTGAGGGTGAAGCACTGTT 2100
QY 2053 CAGCGTGTCAACTACGAGCGCGCGCGCCGCTCTGCTGAGGGTGAAGCACTGTT 2112
Db 2101 CTTGGAGGATATCCACAGGCGCTGGCGCACCTTCTGCTGCTGTGTGCGGCGCCAGACCC 2160
QY 2113 CTTGGAGGATATCCACAGGCGCTGGCGCACCTTCTGCTGCTGTGTGCGGCGCCAGACCC 2172
Db 2161 GCGCGCTGAGCTGCTTGTCAAGTGGATGTGAGGGGCGCTGACGACCACTTCCCCA 2220
QY 2173 GCGCGCTGAGCTGCTTGTCAAGTGGATGTGAGGGGCGCTGACGACCACTTCCCCA 2232
Db 2221 GGACAGGCTCACGAGGCTCATCCAGCATCATCAAAACCCAGAACACGCTGCTGCTGCG 2280
QY 2233 GGACAGGCTCACGAGGCTCATCCAGCATCATCAAAACCCAGAACACGCTGCTGCTGCG 2292
Db 2281 TCGGTATGCGGTGGTCCAGAGCGCCCGCATGGGCGACGCTCCGCAAGCGCTTCAAGAGCCA 2340
QY 2293 TCGGTATGCGGTGGTCCAGAGCGCCCGCATGGGCGACGCTCCGCAAGCGCTTCAAGAGCCA 2352
Db 2341 CGTC----- 2344
QY 2353 CGTCTTACTTTGACAGACCTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTGCAGGA 2412
Db 2344 ----- 2344
QY 2413 GACCAGCGCGCTGAGGGATCCCGTCTCATCGACAGAGCTCTCTCTGATGAGGCCAG 2472
Db 2344 ----- 2344
QY 2473 CAGTGGCTCTTTCGAGCTTCTTCTACGCTTTCATGTGCCACACGCGCTGGGCATCAGGG 2532
Db 2345 ----- 2345
QY 2533 CAACTCTAGCTGCTGCGAGGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2592
Db 2399 CAGCTGTGTCTAGCGGACATGGAGAACAGCTTTTGGGGGATTCGGGGGACGCGCT 2458
QY 2593 CAGCTGTGTCTAGCGGACATGGAGAACAGCTTTTGGGGGATTCGGGGGACGCGCT 2652
Db 2459 GCTCTGCGTTTGGTGGATGATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2518
QY 2653 GCTCTGCGTTTGGTGGATGATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2712
Db 2519 CTTTCTCAGAGCCCTGGTTCGAGGTGTCCCTGATGATGCTGGTGGTGAACCTTGGGAA 2578
QY 2713 CTTTCTCAGAGCCCTGGTTCGAGGTGTCCCTGATGATGCTGGTGGTGAACCTTGGGAA 2772
Db 2579 GACAGTGGTGAACCTTCCCTGTAGAAGACGAGCGCTGGGTGGGCGCTTCTTCTCAGAT 2638
QY 2773 GACAGTGGTGAACCTTCCCTGTAGAAGACGAGCGCTGGGTGGGCGCTTCTTCTCAGAT 2832

Db 2639 GCGGGCCACGGCTATTCCCTGGTGGCGCTGCTGATACCCGGACCTGGAGGT 2698
QY 2833 GCGGGCCACGGCTATTCCCTGGTGGCGCTGCTGATACCCGGACCTGGAGGT 2892
Db 2699 GCAGAGGACTACTCCAGCTATGCCGGAGCTCCATCAGAGCCAGTCTCACTTCAACCG 2758
QY 2893 GCAGAGGACTACTCCAGCTATGCCGGAGCTCCATCAGAGCCAGTCTCACTTCAACCG 2952
Db 2759 GCGCTTCAAGGCTGGGAGAACATGCTCGCAAACTCTTTGGGGTCTTGGGGTGAAGTG 2818
QY 2953 GCGCTTCAAGGCTGGGAGAACATGCTCGCAAACTCTTTGGGGTCTTGGGGTGAAGTG 3012
Db 2819 TCACAGGCTGTTCTGATTTGAGGTGAACAGCTCCAGAGCTGTCACCAACATCTA 2878
QY 3013 TCACAGGCTGTTCTGATTTGAGGTGAACAGCTCCAGAGCTGTCACCAACATCTA 3072
Db 2879 CAAGATCCTCTGCTGAGGCTACAGGTTTCAGCATGTGTGTCAGCTCCCATTTCA 2938
QY 3073 CAAGATCCTCTGCTGAGGCTACAGGTTTCAGCATGTGTGTCAGCTCCCATTTCA 3132
Db 2939 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGGCTCATCTCTGACAGGCGCTCCCT 2998
QY 3133 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGGCTCATCTCTGACAGGCGCTCCCT 3192
Db 2999 CTGCTACTCATCTGAAAGCCAGAGACGAGGATGCTGCGGGCCAGGCGCGC 3058
QY 3193 CTGCTACTCATCTGAAAGCCAGAGACGAGGATGCTGCGGGCCAGGCGCGC 3252
Db 3059 CGGCGCTCTGCTGAGGCGCTGAGTGGCTGTGCAACAGCATTTCTGCTCAAGCT 3118
QY 3253 CGGCGCTCTGCTGAGGCGCTGAGTGGCTGTGCAACAGCATTTCTGCTCAAGCT 3312
Db 3119 GACTGACACCGTGTACCTAGCTGCACTCTGCGGCTGCTGAGGAGACCCAGAGCA 3178
QY 3313 GACTGACACCGTGTACCTAGCTGCACTCTGCGGCTGCTGAGGAGACCCAGAGCA 3372
Db 3179 GCTGAGTGGAGTCTCCGGGAGAGCTGACTGCTGGAGGCGGACCCACCGGC 3238
QY 3373 GCTGAGTGGAGTCTCCGGGAGAGCTGACTGCTGGAGGCGGACCCACCGGC 3432
Db 3239 ACTGCCCTCAGACTTCAAGACCATCTGAGTCTGATGCGCCACCGCCACAGGCGCA 3298
QY 3433 ACTGCCCTCAGACTTCAAGACCATCTGAGTCTGATGCGCCACCGCCACAGGCGCA 3492
Db 3299 GAGCAGACACAGAGCCCTGTACGCGCGGCTTACGTCCTCAGGAGGAGGCGGCGC 3358
QY 3493 GAGCAGACACAGAGCCCTGTACGCGCGGCTTACGTCCTCAGGAGGAGGCGGCGC 3552
Db 3359 CACACCCAGGCGCGCAGCTGGAGTCTGAGGCTGAGTGTGTCGCGAGGCGCTG 3418
QY 3553 CACACCCAGGCGCGCAGCTGGAGTCTGAGGCTGAGTGTGTCGCGAGGCGCTG 3512
Db 3419 CATGTCCGGCTGAGGCTGAGTGTCCGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3478
QY 3613 CATGTCCGGCTGAGGCTGAGTGTCCGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3672
Db 3479 GAGTGTCCAGACACCTGCGCTTCACTTCCCGACAGGCTGCGCTGCGCTCCACCCCA 3538
QY 3673 GAGTGTCCAGACACCTGCGCTTCACTTCCCGACAGGCTGCGCTGCGCTCCACCCCA 3732
Db 3539 GGCGAGCTTTTCTCACCAGGAGCGGCTTCCACTCCCGACATAGGATAGTCCATCC 3598
QY 3733 GGCGAGCTTTTCTCACCAGGAGCGGCTTCCACTCCCGACATAGGATAGTCCATCC 3792
Db 3599 CCAGAT 3604
QY 3793 CCTGAT 3798

RESULT 11
ID US-08-951-733-13 STANDARD; DNA; UNC; 2848 BP.
AC xxxxxx
DT

DE Sequence 13, Application US/08951733
CC Sequence 13, Application US/08951733
CC GENERAL INFORMATION:
CC APPLICANT: Harrington, Lea A.
CC APPLICANT: Robinson, Murray O.
CC TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS
CC NUMBER OF SEQUENCES: 44
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Amgen Inc.
CC STREET: One Amgen Center Drive
CC CITY: Thousand Oaks
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 91320-1789
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent in Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/951,733
CC FILING DATE: 16-OCT-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/873,039
CC FILING DATE: 11-JUN-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/751,189
CC FILING DATE: 15-NOV-1996
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Oleski, Nancy A.
CC REGISTRATION NUMBER: 34,688
CC REFERENCE/DOCKET NUMBER: A-433B
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (805) 447-6504
CC TELEFAX: (805) 499-8011
CC INFORMATION FOR SEQ ID NO: 13:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2848 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC SEQUENCE 2848 BP; 437 A; 978 C; 945 G; 488 T; 0 OTHER.

Query Match 75.0%; Score 2848; DB 22; Length 2848;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2848; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 CAGCGCTCGGGCAGCGCTGCTGCTGCGACGCTGGAGACCCCTGCGCGGCCAC 60
QY 2 CAGCGCTCGGGCAGCGCTGCTGCTGCGACGCTGGAGACCCCTGCGCGGCCAC 61
Db 61 CCGCGCATCCCGCGGCTCCCGCTGCCAGCGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 62 CCGCGCATCCCGCGGCTCCCGCTGCCAGCGCTGCTGCTGCTGCTGCTGCTGCT 121
Db 121 CCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 122 CCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
Db 181 GTGACAGCGGGGACCCCGCGGCTTTCCCGCGGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 182 GTGACAGCGGGGACCCCGCGGCTTTCCCGCGGCTGCTGCTGCTGCTGCTGCTGCT 241
Db 241 CCCTGGAGCAGCG 300
QY 242 CCCTGGAGCAGCG 301
Db 301 GAGCTGCTGGCG 360
QY 302 GAGCTGCTGGCG 361

QY 2522 CGCATCAGGGCAAGTCTACGTCCAGTGCAGGGATCCCGCAGGGCTCCATCCTCTCC 2581
Db 2581 AGCGTGCTCTGCAGCTGTGTACGGGACATGGAGACAGCTGTTTGGGGGATTCCG 2640
QY 2582 AGCGTGCTCTGCAGCTGTGTACGGGACATGGAGACAGCTGTTTGGGGGATTCCG 2641
Db 2641 CGGGACGGGCTGCTGCTGGTTTGTGGATGATTCTTGTGGTGACACCTCACCTCACC 2700
QY 2642 CGGGACGGGCTGCTGCTGGTTTGTGGATGATTCTTGTGGTGACACCTCACCTCACC 2701
Db 2701 CACGGGAAACCTTCTCAGGACCTGTGTCCAGGTGTCCCTGTAGTATGGTTCGCTGGTG 2760
QY 2702 CACGGGAAACCTTCTCAGGACCTGTGTCCAGGTGTCCCTGTAGTATGGTTCGCTGGTG 2761
Db 2761 AACTTGGCGAGACAGTGTGTAACCTCCCTGTAGAACAGAGGCGCTGGTGGACAGGCT 2820
QY 2762 AACTTGGCGAGACAGTGTGTAACCTCCCTGTAGAACAGAGGCGCTGGTGGACAGGCT 2821
Db 2821 TTTGTTTCAGATCCGCGCCACGGCCTAT 2848
QY 2822 TTTGTTTCAGATCCGCGCCACGGCCTAT 2849

RESULT 12
ID US-08-912-951-3 STANDARD; DNA; UNC; 2176 BP.
AC xxxxxx
DT
DE Sequence 3, Application US/08912951
CC Sequence 3, Application US/08912951
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC TITLE OF INVENTION: THERAPEUTIC METHODS
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002600US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2176 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: -
CC LOCATION: 1..2176
CC OTHER INFORMATION: /note= "clone 712562"
CC SQ SEQUENCE 2176 BP; 432 A; 678 C; 642 G; 422 T; 2 OTHER.

Query Match 45.1%; Score 1711; DB 22; Length 2176;
Best Local Similarity 91.2%; Pred. No. 0.00e+00; Mismatches 3; Indels 182; Gaps 1;
Matches 1924; Conservative 0;
Db 1 GGCCAAGTTCCTGCACCTGGCTGATGAGTGTAGTGTGCTGCTGAGTGTGCTCAGGTCTTTCTT 60
QY 1690 GGCCAAGTTCCTGCATGGCTGATGAGTGTGCTGCTGAGTGTGCTGAGTGTGCTCAGGTCTTTCTT 1749
Db 61 TTATGTCAAGGAGACACAGCTTTTCAAAAGAACAGAGCTCTTTTCTACCGGAAGAGTGTCTG 120
QY 1750 TTATGTCAAGGAGACACAGCTTTTCAAAAGAACAGAGCTCTTTTCTACCGGAAGAGTGTCTG 1809
Db 121 GAGCAAGTTGCAAGAGCATTGAATCAGACAGCATTTCAAGAGGGTGCAGCTGCGGGAGCT 180
QY 1810 GAGCAAGTTGCAAGAGCATTGAATCAGACAGCATTTGAAGAGGGTGCAGCTGCGGGAGCT 1809
Db 181 GTCGGAAGCAGAGGTCTCAGGAGCATTCGGGAAGCAGAGCGCCGCCCTCTGCTGACGTCCAGACT 240
QY 1870 GTCGGAAGCAGAGGTCTCAGGAGCATTCGGGAAGCAGAGCGCCGCCCTCTGCTGACGTCCAGACT 1929
Db 241 CCGCTTCATCCCAAGCCTGACGGGCTGCGGGCTGCGGCCGATTGTGAACATGAGTACGTCTGTTGG 300
QY 1930 CCGCTTCATCCCAAGCCTGACGGGCTGCGGGCTGCGGCCGATTGTGAACATGAGTACGTCTGTTGG 1989
Db 301 AGCCAGAACGTTCCGAGAGAAAGAGGGCGGAGCGCTCTCACCTCGAGGGTGAAGGCACCT 360
QY 1990 AGCCAGAACGTTCCGAGAGAAAGAGGGCGGAGCGCTCTCACCTCGAGGGTGAAGGCACCT 2049
Db 361 GTTCAGCGTGTCAACTACGAGCGGGCGCGGCCCTCTCTGGCGGCTCTGTGCT 420
QY 2050 GTTCAGCGTGTCAACTACGAGCGGGCGCGGCCCTCTCTGGCGGCTCTGTGCT 2109
Db 421 GGGCCTGGACGATATCCACAGGGCCTGGCGACCTTCTGCTGCTGCGTGTGCGGGCCCAAGGA 480
QY 2110 GGGCCTGGACGATATCCACAGGGCCTGGCGACCTTCTGCTGCTGCGTGTGCGGGCCCAAGGA 2169
Db 481 CCGCCCGCCTGAGTGTGACTTTGTCAAGGTGGATGTGACGGGCGCGGTACGACACCATCC 540
QY 2170 CCGCCCGCCTGAGTGTGACTTTGTCAAGGTGGATGTGACGGGCGCGGTACGACACCATCC 2229
Db 541 CCAGGACAGGCTCAGGAGGTCTATCGGCAGCATCATCAAAACCCAGAACACGTACTCGCT 600
QY 2230 CCAGGACAGGCTCAGGAGGTCTATCGGCAGCATCATCAAAACCCAGAACACGTACTCGCT 2289
Db 601 GCGTCGATGCGGTGTGTCCAGAGGCGGCCCATGGGCACGTCCGCAAGGCGCTTCAAGAG 660
QY 2290 GCGTCGATGCGGTGTGTCCAGAGGCGGCCCATGGGCACGTCCGCAAGGCGCTTCAAGAG 2349
Db 661 CCAGTGC----- 667
|||||

QY 2350 CCAGCTCTTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCA 2409
Db 667 ----- 667
QY 2410 GGAGACCAGCCCGCTGAGGGATGCCGTGTCATCGAGCAGAGCTCCCTCCCTGAATGAGGC 2459
Db 667 ----- 667
QY 2470 CAGCAGTGGCCCTTTCGACGCTTCCTCAGCTTTCATGTGCCACACGCGCTGCGCATCAG 2529
Db 668 -----CTACGTCAGTCCAGGGGATCCCGCAGGGCTCCCATCTCTCCACGGCTGCT 718
QY 2530 GGGCAAGTCTACGTCAGTCCAGGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCT 2589
Db 719 CTGACGCTGTGTCAGCGCGACATGGAGAACAGCTGTTCGGGGGATTCGGCGGGACGG 778
QY 2590 CTGACGCTGTGTCAGCGCGACATGGAGAACAGCTGTTCGGGGGATTCGGCGGGACGG 2649
Db 779 GCTGCTCTGCTGTTGGTGGATTTCTTGTGGTGACACCTCACCTCACCCACGGAA 838
QY 2650 GCTGCTCTGCTGTTGGTGGATTTCTTGTGGTGACACCTCACCTCACCCACGGAA 2709
Db 839 AACCTTCTCAGGACCTTGGTCCGAGGTGTCCTCAGTATGCTGCGTGGTGAACCTTGGC 898
QY 2710 AACCTTCTCAGGACCTTGGTCCGAGGTGTCCTCAGTATGCTGCGTGGTGAACCTTGGC 2769
Db 899 GAAGACAGTGTGAACCTTCCCTGTAGAGACGAGGCCCTGGGTGGCAGCGCTTTGTTC 958
QY 2770 GAAGACAGTGTGAACCTTCCCTGTAGAGACGAGGCCCTGGGTGGCAGCGCTTTGTTC 2829
Db 959 GATCGCGCCACAGCCCTATTCCCTGTGTGGGCTGCTGTGTATACCCGGACCCCTGGA 1018
QY 2830 GATCGCGCCACAGCCCTATTCCCTGTGTGGGCTGCTGTGTATACCCGGACCCCTGGA 2889
Db 1019 GGTGACAGGACCTACTCAGCTATGCCCGGACCTCCATCAGAGCAGCTCACCCTTCAA 1078
QY 2890 GGTGACAGGACCTACTCAGCTATGCCCGGACCTCCATCAGAGCAGCTCACCCTTCAA 2949
Db 1079 CCGCGGCTTCAAGGCTGGGAGAACATCGCTGCGAAACTCTTTGGGGTCTTTCGGCTGAA 1138
QY 2950 CCGCGGCTTCAAGGCTGGGAGAACATCGCTGCGAAACTCTTTGGGGTCTTTCGGCTGAA 3009
Db 1139 GTGTCACAGCTGTTCTGGATTTGCAGGTGAACAGCCCTCAGAGCGTGTGCACCAACAT 1198
QY 3010 GTGTCACAGCTGTTCTGGATTTGCAGGTGAACAGCCCTCAGAGCGTGTGCACCAACAT 3069
Db 1199 CTACAAGATCCTCTGCTGCGAGGCTACAGTTTCACGCACTGTGCTGCGAGCTCCCAAT 1258
QY 3070 CTACAAGATCCTCTGCTGCGAGGCTACAGTTTCACGCACTGTGCTGCGAGCTCCCAAT 3129
Db 1259 TCATCAGCAAGTTTGGAGAACCCCAACATTTTCTGCGCGTCTCTCTGACAGCGCTC 1318
QY 3130 TCATCAGCAAGTTTGGAGAACCCCAACATTTTCTGCGCGTCTCTCTGACAGCGCTC 3189
Db 1319 CCTGTGCTACTCTCTGTAAGCCCAAGCAGGAGTGTGCTGGGGGCGCAGGGGCG 1378
QY 3190 CCTGTGCTACTCTCTGTAAGCCCAAGCAGGAGTGTGCTGGGGGCGCAGGGGCG 3249
Db 1379 CGCGGGCCCTGCGCTCCGAGGCGGTGCACTGCGTGTGCGCAAGCAATCTCTGCTCAA 1438
QY 3250 CGCGGGCCCTGCGCTCCGAGGCGGTGCACTGCGTGTGCGCAAGCAATCTCTGCTCAA 3309
Db 1439 GCTGACTCGACACCGTGTACCTAGTGTGCGTCTCTGCGGCTCACTCAGGACGCGCCAGAC 1498
QY 3310 GCTGACTCGACACCGTGTACCTAGTGTGCGTCTCTGCGGCTCACTCAGGACGCGCCAGAC 3369
Db 1499 GCAGTGTGAGTGTGAGGAGTCCCGGGGACGAGCTGACTGCTGAGGCGCGGACCAACC 1558
QY 3370 GCAGTGTGAGTGTGAGGAGTCCCGGGGACGAGCTGACTGCTGAGGCGCGGACCAACC 3429
Db 1559 GGCAGTGTGCTCAGACTTCAAGACCATCTTGACTGATGGCCACCCCGCCACGCGAGGC 1618
QY 3430 GGCAGTGTGCTCAGACTTCAAGACCATCTTGACTGATGGCCACCCCGCCACGCGAGGC 3489

RESULT 13

ID US-09-026-981-36 STANDARD; DNA; UNC; 3346 BP.
AC xxxxxx

DE Sequence 36, Application US/09026981
Sequence 36, Application US/09026981
GENERAL INFORMATION:
CC APPLICANT: Counter, Christopher M.
CC APPLICANT: Meyerson, Matthew
CC APPLICANT: Weinberg, Robert A.
CC TITLE OF INVENTION: telomerase Catalytic Subunit Gene and
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Militia Drive
CC CITY: Lexington
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/026,981
CC FILING DATE: 20-FEB-1998
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/064,322
CC FILING DATE: 30-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/055,762
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/054,549
CC FILING DATE: 01-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/047,151
CC FILING DATE: 20-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/038,750
CC FILING DATE: 20-FEB-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WHI97-11p4M

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 781-861-6240
CC TELEFAX: 781-861-9540
CC INFORMATION FOR SEQ ID NO: 36:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 3346 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
SQ SEQUENCE 3346 BP: 617 A; 1027 C; 965 G; 735 T; 2 OTHER.

Query Match 41.5%; Score 1575; DB 23; Length 3346;
Best Local Similarity 85.2%; Pred. No. 0.00e+00;
Matches 1975; Conservative 0; Mismatches 1; Indels 341; Gaps 2;

Db 971 GGGTTGGCTGTTCCTCCGCGCAGAGCACCGTCTCGGTGAGGAGATCCTGGCCAAAGTTCC 1030
Qy 1641 GGGTTGGCTGTTCCTCCGCGCAGAGCACCGTCTCGGTGAGGAGATCCTGGCCAAAGTTCC 1700
Db 1031 TGCACCTGGCTGATGATGTGTACGTGTCGAGCTGCTCAGGTCTCTTTTATGTCACGG 1090
Qy 1701 TGCACCTGGCTGATGATGTGTACGTGTCGAGCTGCTCAGGTCTCTTTTATGTCACGG 1760
Db 1091 AGACCACGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAGATTGC 1150
Qy 1761 AGACCACGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAGATTGC 1820
Db 1151 AAAGCATTGGATCAGACAGCACTTGAAGAGGGTGCAGCTCGGGAGGTGTCGGAAGCAG 1210
Qy 1821 AAAGCATTGGATCAGACAGCACTTGAAGAGGGTGCAGCTCGGGAGGTGTCGGAAGCAG 1880
Db 1211 AGGTGAGCAGCATCGGAAGCAGGCGCGCTGCTGACGTCCAGATCTCGCTTCATCC 1270
Qy 1881 AGGTGAGCAGCATCGGAAGCAGGCGCGCTGCTGACGTCCAGATCTCGCTTCATCC 1940
Db 1271 CCAAGCTTGACGGCTGGCGCGATTTGACATGAGTGTCTGAGGAGCAGAAAGT 1330
Qy 1941 CCAAGCTTGACGGCTGGCGCGATTTGACATGAGTGTCTGAGGAGCAGAAAGT 2000
Db 1331 TCCGCAGAGAAAGAGGCGCAGCGTCTCACCTCGAGGGTGAAGCACTGTTACGCGTGC 1390
Qy 2001 TCCGCAGAGAAAGAGGCGCAGCGTCTCACCTCGAGGGTGAAGCACTGTTACGCGTGC 2060
Db 1391 TCAACTAGAGGGCGGGCGCGCGCTGCTGCGGCGCTCTGTGCTGGGCGTGGAGC 1450
Qy 2061 TCAACTAGAGGGCGGGCGCGCGCTGCTGCGGCGCTCTGTGCTGGGCGTGGAGC 2120
Db 1451 ATATCCAGAGGCGTGGCGCACCTTCGTGCTGCGGTGCGGGCCAGACCGCGCGCTG 1510
Qy 2121 ATATCCAGAGGCGTGGCGCACCTTCGTGCTGCGGTGCGGGCCAGACCGCGCGCTG 2180
Db 1511 AGCTGTACTTTTCAAGGTGGATGTGACGGGCGGTAGCAGACCATCCCCCAGGACAGGC 1570
Qy 2181 AGCTGTACTTTTCAAGGTGGATGTGACGGGCGGTAGCAGACCATCCCCCAGGACAGGC 2240
Db 1571 TCACGGAGGTCTATCCAGCATCATCAAAACCCAGAACACGTACTGTGCGTGGGTATG 1630
Qy 2241 TCACGGAGGTCTATCCAGCATCATCAAAACCCAGAACACGTACTGTGCGTGGGTATG 2300
Db 1631 CGGTGCTCCAGAGGCGCGCATGGCAGCTCCGCAAGGCTTCAAGAGCCACGTC 1686
Qy 2301 CGGTGCTCCAGAGGCGCGCATGGCAGCTCCGCAAGGCTTCAAGAGCCACGTCCTA 2360
Db 1686 ----- 1686
Qy 2361 CTTTGACAGACCTCCAGCGGTACATGCGACAGTTGCTGGCTCACCTGCGAGGACCGCC 2420
Db 1686 ----- 1686
Qy 2421 CGCTGAGGGATGCCGTCTATCGAGCAGAGCTCTCCCTGAATGAGGCCAGCAGTGCC 2480
Db 1687 -----CT 1688
||

Qy 2481 TCCTCGAGCTCTTCTACGCTTCAATGTGCCACACGCGCGTGGCATCAGGGGCAAGTCCT 2540
Db 1689 ACGTCCAGTCCAGGGGATCCCGAGGCTCCATCTCTCCACGCTGCTCTGACGCTGT 1748
Qy 2541 ACGTCCAGTCCAGGGGATCCCGAGGCTCCATCTCTCCACGCTGCTCTGACGCTGT 2600
Db 1749 GCTACGGGACATGGAGAAACAGCTGTTTGGGGGATTCGGCGGAGCGGGTGTCTCTGC 1808
Qy 2601 GCTACGGGACATGGAGAAACAGCTGTTTGGGGGATTCGGCGGAGCGGGTGTCTCTGC 2660
Db 1809 GTTTGGTGGATGATTCTTGTGGTGACACTCACTCACTCACTCACTCACTCACTCA 1868
Qy 2661 GTTTGGTGGATGATTCTTGTGGTGACACTCACTCACTCACTCACTCACTCACTCA 2720
Db 1869 GGACCTCGTCCGAGGTGCTCCCTGAGTATGGCTGCTGGTGAACCTTCGGGAAGACATGG 1928
Qy 2721 GGACCTCGTCCGAGGTGCTCCCTGAGTATGGCTGCTGGTGAACCTTCGGGAAGACATGG 2780
Db 1929 TGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGSCACGGCTTTTGTTCAGATGCCGGCC 1988
Qy 2781 TGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGSCACGGCTTTTGTTCAGATGCCGGCC 2840
Db 1989 ACGSCCTATTCCCTGCTGGTGGCTGCTGCTGGATACCCGAGCCCTGGAGGTGCAGAGC 2048
Qy 2841 ACGSCCTATTCCCTGCTGGTGGCTGCTGCTGGATACCCGAGCCCTGGAGGTGCAGAGC 2900
Db 2049 ACTACTCCAGTATGCCCGGACCTTCCATCAGAGCCAGTCTCACCTTCAACCGCGGCTTCA 2108
Qy 2901 ACTACTCCAGTATGCCCGGACCTTCCATCAGAGCCAGTCTCACCTTCAACCGCGGCTTCA 2960
Db 2109 AGGTGAGGAGACATCGCTCGCAAACTCTTTGGGGTCTTGGGGCTGAAGTGTACAGCC 2168
Qy 2961 AGGTGAGGAGACATCGCTCGCAAACTCTTTGGGGTCTTGGGGCTGAAGTGTACAGCC 3020
Db 2169 TGTTCCTGGATTTGACAGTGAACAGCCTCCAGAGGTGTGCACCAACATCTACAGATCC 2228
Qy 3021 TGTTCCTGGATTTGACAGTGAACAGCCTCCAGAGGTGTGCACCAACATCTACAGATCC 3080
Db 2229 TCCTGCTGACGGGTACAGGTTTCACGCATGTGTGCTGACGCTCCCATTTTCATCAGCAAG 2288
Qy 3081 TCCTGCTGACGGGTACAGGTTTCACGCATGTGTGCTGACGCTCCCATTTTCATCAGCAAG 3140
Db 2289 TTGGAAGAACCCACATTTTCTGCGGTCTATCTGTGACAGGCTCTCCCTCTGCTACT 2348
Qy 3141 TTGGAAGAACCCACATTTTCTGCGGTCTATCTGTGACAGGCTCTCCCTCTGCTACT 3200
Db 2349 CCATCTCTGAAGCCAAACAGCAGCCGCAAGAAACATTTCTGTGCTGACTCTCTCGGCTGC 2408
Qy 3201 CCATCTCTGAAGCCAAACAGCAG - - - - - 3224
Db 2409 TTGGTCCGGACAGCCAGAGATGGAGCCACCCCGCAGACCGTGGGTGTGGGCAAGCTTC 2468
Qy 3224 ----- 3224
Db 2469 CGGTGCTCTCTGGAGGGGAGCTGGGGCTGGGCTGTGACTCTCTCAGCTCTGTGTTTCCCC 2528
Qy 3224 ----- 3224
Db 2529 CAGGATGTCGCTGGGGCCAAAGGGCGCGCGCGCTCTGCGCTCCGAGGCGCGTGCAGT 2588
Qy 3225 ---GGATGTGCTGGGGCCAAAGGGCGCGCGCGCTCTGCGCTCCGAGGCGCGTGCAGT 3281
Db 2589 GGTGTGCCACCAAGCATTTCTGCTCAAGTGTACTGACACCGCTGTCACTACGTGCTGCAC 2648
Qy 3282 GGTGTGCCACCAAGCATTTCTGCTCAAGTGTACTGACACCGCTGTCACTACGTGCTGCAC 3341
Db 2649 TCCTGGGGTCACTCAGGACAGCCAGCAGCTGAGTTCGGAAGCTCCCGGGGACGACG 2708
Qy 3342 TCCTGGGGTCACTCAGGACAGCCAGCAGCTGAGTTCGGAAGCTCCCGGGGACGACG 3401
Db 2709 TGACTGCCCTGGAGGGCGCAGCCAAACCGCACTGCCCTCAGACTTCAAGACCATCTG 2768
Qy 3402 TGACTGCCCTGGAGGGCGCAGCCAAACCGCACTGCCCTCAGACTTCAAGACCATCTG 3461

QY 2411 GAGACACCGCGCTGAGGATGCCGTGCTCATCAGCAGAGCTCCTCCCTGAATGAGCC 2470
Db 665 ----- 665
QY 2471 AGCAGTGGCCTCTTCGACGCTTCTCCTACGCTTCATGTGCCACACGCGTGGCATCAGG 2530
Db 666 -----CTACGTCACAGTGCACAGGGATCCCGAGGGCTCCATCTCTCCACGCTGCTC 717
QY 2531 GGCAAGTCCCTACGTCACAGTGCACAGGGATCCCGAGGGCTCCATCTCTCCACGCTGCTC 2590
Db 718 TGACAGCTGTGCTACGCGGACATGAGAACAAAGCTGTTTGGGGGATTCGCGGGAGCGG 777
QY 2591 TGACAGCTGTGCTACGCGGACATGAGAACAAAGCTGTTTGGGGGATTCGCGGGAGCGG 2650
Db 778 CTGCTCTCGGTTTGGTGGATGATTTCTTGTGTGACACCTCACCTACCCAGCGGAAA 837
QY 2651 CTGCTCTCGGTTTGGTGGATGATTTCTTGTGTGACACCTCACCTACCCAGCGGAAA 2710
Db 838 ACCTTCCTCAGGACCTGCTCCGAGGTGTCCTGAGTATGGCTGCGTGAAGTTCGGG 897
QY 2711 ACCTTCCTCAGGACCTGCTCCGAGGTGTCCTGAGTATGGCTGCGTGAAGTTCGGG 2770
Db 898 AAGACAGTGTGAACTTCCCTGTAGAGAGAGGCGCTGGGTGGCAGCGCTTTTGTTCAG 957
QY 2771 AAGACAGTGTGAACTTCCCTGTAGAGAGAGGCGCTGGGTGGCAGCGCTTTTGTTCAG 2830
Db 958 ATGCGCGCCACAGCGCTTATTCCTGTGTGGGCTGCTGTGATACCCGAGCCCTGGAG 1017
QY 2831 ATGCGCGCCACAGCGCTTATTCCTGTGTGGGCTGCTGTGATACCCGAGCCCTGGAG 2890
Db 1018 GTGACAGCGGACTACTCCAGTATGCCGAGACCTCCATCAGAGCGAGTCTACCTTCAAC 1077
QY 2891 GTGACAGCGGACTACTCCAGTATGCCGAGACCTCCATCAGAGCGAGTCTACCTTCAAC 2950
Db 1078 CGCGGCTTAAAGGCTGGGAGGAACATCGTCGCAAACTTTTGGGGTCTTGGGCTGAAG 1137
QY 2951 CGCGGCTTAAAGGCTGGGAGGAACATCGTCGCAAACTTTTGGGGTCTTGGGCTGAAG 3010
Db 1138 TGTACAGCGCTGTTTCTGGATTTGAGGTGAACAGCGCTCCAGAGCGGTGTCACCAACATC 1197
QY 3011 TGTACAGCGCTGTTTCTGGATTTGAGGTGAACAGCGCTCCAGAGCGGTGTCACCAACATC 3070
Db 1198 TACAGATCTCTGCTGCAGGCGTACAGGTTTACAGCATGTGTGTCAGAGTCCCATTT 1257
QY 3071 TACAGATCTCTGCTGCAGGCGTACAGGTTTACAGCATGTGTGTCAGAGTCCCATTT 3130
Db 1258 CATCAGCAAGTTTGGAAAGAACCC-ACATTTTCTCGCGCTCATCTCTGACAGGC-TCC 1315
QY 3131 CATCAGCAAGTTTGGAAAGAACCCACATTTTCTCGCGCTCATCTCTGACAGGCCTCC 3190
Db 1316 CTCTGTACTCCATCTCTGAAAGCAAGAACGAGGGATGCTGTTGGGGGCAAGGGCGCC 1375
QY 3191 CTCTGTACTCCATCTCTGAAAGCAAGAACGAGGGATGCTGTTGGGGGCAAGGGCGCC 3250
Db 1376 GCGGGCCTCTGCGCTCCGAGGCGTGCAGTGGTGTGCCACCAAGCATTCCTGCTCAAG 1435
QY 3251 GCGGGCCTCTGCGCTCCGAGGCGTGCAGTGGTGTGCCACCAAGCATTCCTGCTCAAG 3310
Db 1436 CTGACTCGACACCGTGTCACTAGTGCACCTCTCGGGTCACTCAGGACAGCCAGACG 1495
QY 3311 CTGACTCGACACCGTGTCACTAGTGCACCTCTCGGGTCACTCAGGACAGCCAGACG 3370
Db 1496 CAGTGTAGTCGGAAGTCTCCGGGACAGCGTCACTGCCCTGGAGGCGCGAGCAACCGG 1555
QY 3371 CAGTGTAGTCGGAAGTCTCCGGGACAGCGTCACTGCCCTGGAGGCGCGAGCAACCGG 3430
Db 1556 GCAGTGCCTCAGACTTCAAGACCATCCTGTGACTGATGGCCACCGCCACAGCGAGCC 1615
QY 3431 GCAGTGCCTCAGACTTCAAGACCATCCTGTGACTGATGGCCACCGCCACAGCGAGCC 3490
Db 1616 GAGACGACACACGAGCGCTGTGACGCGGGCTTATAGTCTCCAGGGAGGGGGCG 1675
QY 3491 GAGACGACACACGAGCGCTGTGACGCGGGCTC-TACGTCCCGAGGAGGGGGCG 3549

Db 1676 GCCACACCCAGCGCTGCACCGCTGGAGTCTGAGGCTTGAAGTGTGTTGGCCGAGGC 1735
QY 3550 GCCACACCCAGCGCGCCGCTGGAGTCTGAGGCTTGAAGTGTGTTGGCCGAGGC 3609
Db 1736 CTGATGTCCGCTGAAGGCTGAGTGTCCGGCTGAGGCTTGAAGTGTCCAGCCAAAG 1795
QY 3610 CTGATGTCCGCTGAAGGCTGAGTGTCCGGCTGAGGCTTGAAGTGTCCAGCCAAAG 3669
Db 1796 GCTGAGTGTCCAGCACACCTGC-GTTTTTCACTTCCACAGGCTGCGTTCGG-TCCACC 1853
QY 3670 GCTGAGTGTCCAGCACACCTGCGGTCTTCACTTCCACAGGCTGCGTTCGGTCCACC 3729
Db 1854 CCAGGCGCAGCTTTTCTCACCAGGAGCGGCTTCACTTCCCATCCCATAGGAATAGTCCA 1913
QY 3730 CCAGGCGCAGCTTTTCTCACCAGGAGCGGCTTCACTTCCCATAGGAATAGTCCA 3789
Db 1914 TCCCCAGAT 1922
QY 3790 TCCCCAGAT 3798
RESULT 15
ID US-08-954-050-100 STANDARD; DNA; UNC; 2171 BP.
AC xxxxxx
DT
Sequence 100, Application US/08854050
DE Sequence 100, Application US/08854050
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Novel Telomerase
CC NUMBER OF SEQUENCES: 225
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429

CC REFERENCE/DOCKET NUMBER: 015389-002930US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 100:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2171 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 22..1716

CC OTHER INFORMATION: /note= "EcoRI-NotI insert of
CC OTHER INFORMATION: clone 712562 encoding 63 kDa
CC OTHER INFORMATION: telomerase protein"
CC

CC
SQ SEQUENCE 2171 BP: 433 A; 667 C; 641 G; 429 T; 1 OTHER.
Query Match 40.0%; Score 1520; DB 21; Length 2171;
Best Local Similarity 90.8%; Pred. No. 0.00e+00;
Matches 1915; Conservative 1; Mismatches 5; Indels 188; Gaps 7;

Db 1 GCCAAGTTCTCGCACTGCTGATGAGTGTACGTCGTCGAGCTGCTCAGCTCTTCTTT 60
Qy 1691 GCCAAGTTCTCGCACTGCTGATGAGTGTACGTCGTCGAGCTGCTCAGGCTCTTCTTT 1750
Db 61 TATGTCAGCGAGACACCACTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGG 120
Qy 1751 TATGTCAGCGAGACCACTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGG 1810
Db 121 AGCAAGTTGCAAGCATTTGGAATCAGACAGACACTTTGAAGAGGTCAGCTGCGGGAGCTG 180
Qy 1811 AGCAAGTTGCAAGCATTTGGAATCAGACAGACACTTTGAAGAGGTCAGCTGCGGGAGCTG 1870
Db 181 TCGGAAGCAGAGGTCAGCAGCATCGGAAGCCAGCGCCGCTGCTCAGCTCCAGACTC 240
Qy 1871 TCGGAAGCAGAGGTCAGCAGCATCGGAAGCCAGCGCCGCTGCTCAGCTCCAGACTC 1930
Db 241 CGCTTATCCCAAGCCCTGACGGGCTGGGGCGGATTTGAACATGGACTAGTCTGTGGGA 300
Qy 1931 CGCTTATCCCAAGCCCTGACGGGCTGGGGCGGATTTGAACATGGACTAGTCTGTGGGA 1990
Db 301 GCCAAGAGTTCCCGCAGAGAAAGARGCCGAGCGTCTCACTCAGAGGTTGAAGCACTG 360
Qy 1991 GCCAAGAGTTCCCGCAGAGAAAGARGCCGAGCGTCTCACTCAGAGGTTGAAGCACTG 2050
Db 361 TTCAGCGTGTCAACTAGCAGCGGCGCGCGCCGCTCCTGGGCGCTCTGTGTGCTG 420
Qy 2051 TTCAGCGTGTCAACTAGCAGCGGCGCGCGCCGCTCCTGGGCGCTCTGTGTGCTG 2110
Db 421 GGCCTGGACGATATCCAGAGGCTTGGCGCACCTTCGTGCTGCTGCTGGGCGCCAGGAC 480
Qy 2111 GGCCTGGACGATATCCAGAGGCTTGGCGCACCTTCGTGCTGCTGCTGGGCGCCAGGAC 2170
Db 481 CGCGCGCTGAGCTGACTTTGTCAAGTGTGATGTGACGGCGGTACGACACCATCCCC 540
Qy 2171 CGCGCGCTGAGCTGACTTTGTCAAGTGTGATGTGACGGCGGTACGACACCATCCCC 2230
Db 541 CAGGACAGGCTCAGGAGGTATCCGACAGCATCATCAACCCGAGAAACAGTACTGCGTG 600
Qy 2231 CAGGACAGGCTCAGGAGGTATCCGACAGCATCATCAACCCGAGAAACAGTACTGCGTG 2290
Db 601 CQTGCGTATCGGTGTCAGAAAGCCGCC - ATGGGACGCTCCGCAAGGCTTCAAGAGC 659
Qy 2291 CQTGCGTATCGGTGTCAGAAAGCCGCCATGGGACGCTCCGCAAGGCTTCAAGAGC 2350
Db 660 CAGGTC----- 665
Qy 2351 CAGGTCCTTACCTTGACAGACCTCCAGCGGTACATGCGACAGTTGCTGGCTCAGCTGCGAG 2410
Db 665 ----- 665

Qy 2411 GAGACCAGCGCGCTGAGGGATGCCGTGTCATCGACGAGAGCTCCTCCTGATGAGGCC 2470
Db 665 ----- 665
Qy 2471 AGCAGTGGCCTCTTCGAGCTCTTCTACGTTCATGTGCCACCACGCGGTGCGCATCAGG 2530
Db 666 -----CTACGTCCAGTGGCAGGGATCCCGAGGGTCCATCTCTCCACGCTGCTC 717
Qy 2531 GGCAAGTCCCTACGTCCAGTGGCAGGGATCCCGAGGGTCCATCTCTCCACGCTGCTC 2590
Db 718 TGCAGCTGTGTACGCGGACATGAGAACAGCTTTTGGGGGATTCGGGCGGACGGG 777
Qy 2591 TGCAGCTGTGTACGCGGACATGAGAACAGCTTTTGGGGGATTCGGGCGGACGGG 2650
Db 778 CTGCTCTCGCTTTGGTGGATGATTTCTTTGTGTGACACCTCACCTCACCCACGCGAAA 837
Qy 2651 CTGCTCTCGCTTTGGTGGATGATTTCTTTGTGTGACACCTCACCTCACCCACGCGAAA 2710
Db 838 ACCTTCTCAGGACCCCTGGTCCGAGGTGCTCCTGAGTATGGCTCGGTGGTGAACCTTGGGG 897
Qy 2711 ACCTTCTCAGGACCCCTGGTCCGAGGTGCTCCTGAGTATGGCTCGGTGGTGAACCTTGGGG 2770
Db 898 AAGACAGTGTGAACTTCCCTGTAGAACAGAGGCGCTGGTGGCAGCGCTTTTGTTCAG 957
Qy 2771 AAGACAGTGTGAACTTCCCTGTAGAACAGAGGCGCTGGTGGCAGCGCTTTTGTTCAG 2830
Db 958 ATGCGCGCCACGCGCTATTCCCTGGTGGGCGCTGCTGTGGATACCGGACCTGGAG 1017
Qy 2831 ATGCGCGCCACGCGCTATTCCCTGGTGGGCGCTGCTGTGGATACCGGACCTGGAG 2890
Db 1018 GTGACAGAGCACTACTCCAGCTATGCGCGGACCTTCCATCAGAGCCAGTCTCACCTTCAAC 1077
Qy 2891 GTGACAGAGCACTACTCCAGCTATGCGCGGACCTTCCATCAGAGCCAGTCTCACCTTCAAC 2950
Db 1078 GCGGCTTCAAGGCTGGGAGAACATGCGTGGCAAACTCTTTGGGGTCTTGGCGCTGAAG 1137
Qy 2951 GCGGCTTCAAGGCTGGGAGAACATGCGTGGCAAACTCTTTGGGGTCTTGGCGCTGAAG 3010
Db 1138 TGTACAGCGCTGTTCTCGGATTTTCAGGTGAACAGCTCCAGAGCGGTGTCACCAACATC 1197
Qy 3011 TGTACAGCGCTGTTCTCGGATTTTCAGGTGAACAGCTCCAGAGCGGTGTCACCAACATC 3070
Db 1198 TACAAGATCTCTGCTGACGCGCTACAGTTTTCACGATGTGTGCTGACAGTCCCATTT 1257
Qy 3071 TACAAGATCTCTGCTGACGCGCTACAGTTTTCACGATGTGTGCTGACAGTCCCATTT 3130
Db 1258 CATCAGAAAGTTTGGAAAGAACCC - ACATTTTCTGCGCGTCACTCTGACACAGGC - TCC 1315
Qy 3131 CATCAGAAAGTTTGGAAAGAACCCCACTTTTCTGCGCGTCACTCTGACACAGGCCTCC 3190
Db 1316 CTCTGCTACTCCATCTGAAAGCCAAAGAACGAGGATGTCGTGGGGCCCAAGGGCGCC 1375
Qy 3191 CTCTGCTACTCCATCTGAAAGCCAAAGAACGAGGATGTCGTGGGGCCCAAGGGCGCC 3250
Db 1376 GCCGGCCCTGCGCCTCCGAGGCGGTGACGTGTGTGCCACCAAGCATTCCTGCTCAAG 1435
Qy 3251 GCCGGCCCTGCGCCTCCGAGGCGGTGACGTGTGTGCCACCAAGCATTCCTGCTCAAG 3310
Db 1436 CTGACTGCACACCGTGTCACTAGTGCACCTCTCTGGGTCACCTCAGACAGCCAGAGC 1495
Qy 3311 CTGACTGCACACCGTGTCACTAGTGCACCTCTCTGGGTCACCTCAGACAGCCAGAGC 3370
Db 1496 CAGCTGAGTCCGAAAGTCCCGGGGACGCGTACTGCCCTGGAGGCGCGAGCCCAACCCG 1555
Qy 3371 CAGCTGAGTCCGAAAGTCCCGGGGACGCGTACTGCCCTGGAGGCGCGAGCCCAACCCG 3430
Db 1556 GCATGCGCCTCAGACTTCAAGACCATCTGAGCTGTATGGCCACCGCCGCCACAGCCAGGCG 1615
Qy 3431 GCATGCGCCTCAGACTTCAAGACCATCTGAGCTGTATGGCCACCGCCGCCACAGCCAGGCG 3490
Db 1616 GAGAGCAGACACAGAGCCCTGTACGCGGGCTTTATAGTCCAGGAGGAGGGGCG 1675
Qy 3491 GAGAGCAGACACAGAGCCCTGTACGCGGGGCTC - TAGTCCAGGAGGAGGGGCG 3549

Db 1676 GCCCACCCAGCCCTGCACCCGCTGGAGTCTGAGGCCCTGAGTGTGTTGGCCGAGGC 1735
|||||
QY 3550 GCCCACCCAGCCCGCCACCCGCTGGAGTCTGAGGCCCTGAGTGTGTTGGCCGAGGC 3609
|||||
Db 1736 CTGCATGTCGGCTGAAGCTGAGTGTCCGGCTGAGGCCCTGAGCGAGTGTCCAGCCAGG 1795
|||||
QY 3610 CTGCATGTCGGCTGAAGCTGAGTGTCCGGCTGAGGCCCTGAGCGAGTGTCCAGCCAGG 3669
|||||
Db 1796 GGTGAGTGTCCAGCACACCTGC-GTTTTCACTTCCCACAGGCTGGCGTTCCG-TCCACC 1853
|||||
QY 3670 GGTGAGTGTCCAGCACACCTGCCGTCTTCACTTCCCACAGGCTGGCGTCCGCTCCACC 3729
|||||
Db 1854 CCAGGCCAGCTTTTCTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCA 1913
|||||
QY 3730 CCAGGCCAGCTTTTCTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCA 3789
|||||
Db 1914 TCCCCAGAT 1922
|||||
QY 3790 TCCCCGTAT 3798

Search completed: Fri Dec 25 08:54:06 1998
Job time : 8861 secs.